Australian National University

Audit Firm Industry Specialisation and Analyst Forecast Accuracy

Yi Wu

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

June 2014

DECLARATION

I, Yi Wu, hereby declare that this thesis is my original work, and that it contains no material previously published or written by another person, except where due acknowledgement is made in the text.

Two papers drawn from the evidence presented in this thesis have been presented at international academic conferences and are co-authored with my supervisor.

Yi Wu

ACKNOWLEDGEMENTS

Completing a PhD is a long and challenging journey, and I am realising that I am at the end of this journey. I would like to express my thanks to all those who contributed in many ways to the success of this thesis and made it an unforgettable and rewarding experience for me.

Foremost, I owe my deepest gratitude to my supervisor, Associate Professor Mark Wilson, for his continuous support of my PhD study. He has been patient and encouraging, and has challenged me and shared with me his inspiring guidance and invaluable advice in many respects. I would also like to thank him for his understanding and support during the hardest period of my PhD years. I have been extremely fortunate to have him as my supervisor, and without him I would not have been able to complete this doctoral thesis.

I am grateful to my advisors, Associate Professor Greg Shailer and Professor Baljit Sidhu, for their encouragement and useful advice on my work. I also thank Professor Neil Fargher for giving the interesting seminar in advanced auditing that inspired me to undertake research in this area.

I have also been glad and fortunate to study and work with a group of knowledgeable, helpful and friendly fellow PhD students and colleagues at the Australian National University. I thank Professor Juliana Ng and Dr Janet Lee for their helpful career advice and suggestions in general.

I would also like to thank my friends, Louise Lu, Isabel Wang, Libby Zhang, Gladys Lee, Steven Wu, Yangxin Yu, Tejshree Kala, Sonali Walpola, Tracy Wang and Jenny Lee, for providing a stimulating and fun-filled 'corridor' environment. Without their company, the journey would have been lonely.

Last but not least, I am greatly indebted to my parents, Hongtao Wu and Xiumei Tang, and my sister, Jie Wu, for their understanding and unconditional support. I have been fortunate to have met and married my husband, Qing Yu, during my study. When they see me obtain my doctorate, I am certain they will be the happiest and proudest of me. Their support and love have given me the strength to complete this journey.

ABSTRACT

My thesis examines whether the extent to which audit firms concentrate their business in particular industries ('audit firm industry specialisation') improves the usefulness of published financial reports for analysts' predictions of future earnings, and whether the strength of any observed association varies in a manner consistent with the existence of a causal relationship between audit quality and analyst forecast accuracy. Prior research presents diametrically opposite predictions and results regarding the directional relationship between audit firm industry specialisation and analyst forecast accuracy. My thesis shows that the conflicting results in the literature arise largely from prior studies' focus on short-horizon (end-of-year) forecast accuracy, which is subject to competing effects related to audit quality, and which in turn renders the resulting empirical models highly sensitive to model specification. I argue that analyst long-horizon (beginning-of-year) forecast accuracy is a more direct measure of the usefulness of published earnings for the prediction of future performance, and demonstrate that regressions using this metric consistently report a significant positive relation between audit firm industry specialisation and forecast accuracy. I then examine whether the observed positive association between audit firm industry specialisation and forecast accuracy varies with factors argued to reflect the relative importance of audit quality to the predictability of earnings. First, I show that the impact of audit firm industry specialisation on forecast accuracy increases with the underlying riskiness of clients' operations (proxied by cash flow volatility and innate accrual quality). I then argue that audit firm industry specialisation should have a greater impact on the forecast accuracy of lower-quality analysts (where quality is proxied by experience, employer size, 'All-Star' status and

composite measures), who rely relatively heavily on published earnings when generating forecasts. To this end, I present evidence that audit firm industry specialisation has a greater impact on forecast accuracy for: (1) firm-years where the average 'quality' of analysts covering the firm is lower, and (2) for forecasts issued by individual analysts of lower quality. My results are robust to the use of controls for the endogenous selection of industry specialist auditors. In sum, my study presents evidence that greater audit quality does improve the usefulness of financial statements for the prediction of future earnings.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2SLS	Two-stage least squares
CAR	Cumulative abnormal returns
DWH	Durban-Wu-Hausman
ERC	Earnings response coefficient
FRC	Financial Reporting Council
II	Institutional Investor
IOSCO	International Organisation of Securities Commissions
IMR	Inverse *MILLS* ratio
LM	Lagrange multiplier
MSA	Metropolitan statistical area
OLS	Ordinary least squares
PSM	Propensity Score Matching
Reg FD	Regulation Fair Disclosure
ROC	Receiver operating characteristic curve
SAS	Statement on auditing standards
SEC	Securities and Exchange Commission
SFAC	Statement of financial accounting concepts
VIF	Variance inflation factor
WSJ	Wall Street Journal

CHAPTER 1: INTRODUCTION

1.1 Introduction

The purpose of financial reporting is to provide information useful to a maximum number of primary users for the purpose of making decisions about whether to provide resources to the firm (Statement of Financial Accounting Concepts [SFAC] No. 8, FASB 2010 OB8 and OB1). Primary users include existing and potential investors, lenders and other creditors (SFAC No. 8, FASB 2010 OB2 and BC1.9), but it is acknowledged that even well-informed and diligent users may need to seek professional advice, such as that provided by securities analysts, to understand information about complex phenomena (SFAC No. 8, FASB 2010 QC32). Central to users' resource allocation decision is the prediction of firms' ability to generate future cash flows (SFAC No. 8, FASB 2010 OB2 and OB3), critical to which is the information about a firm's past and expected future earnings (SFAC No. 8, FASB 2010 BC1.31). It follows that high-quality financial reports should be useful for predicting firms' future earnings and, *ceteris paribus*, that usefulness should extend to a broad range of financially competent users. My thesis examines whether the extent to which audit firms concentrate their business in particular industries ('audit firm industry specialisation') affects the usefulness (quality) of clients' published financial reports for securities analysts' predictions of future earnings, and whether the strength of any observed relationship varies in a manner consistent with theory regarding the relative influence of audit quality on earnings predictability.

The demand for external financial reporting arises largely in response to agency conflicts originating between self-interested management and equity investors who are unable to observe directly management's behaviour and opportunity set (Jensen and Meckling 1976).[1] One means of mitigating agency costs is the introduction of monitoring or bonding mechanisms designed to align the interests of the principal(s) and agent(s) (Jensen and Meckling 1976). The requirement that management reports to shareholders regarding the use of resources entrusted to them is one of such mechanisms, and the quality of these financial reports is critical to reducing agency costs (Watts and Zimmerman 1978). Mandated accounting standards establish rules governing the preparation of financial reports and constrain the range of acceptable methods for recording and reporting transactions. However, financial reports are prepared by management (the agent), who may exercise discretion over accounting estimates and reporting choices to bias the financial reports in pursuit of their self-interest. This gives rise to a demand for the third-party verification of financial reports and the management assertions underpinning them (Benston 1985).

The external audit service involves the provision of an auditor's opinion about whether the financial statements are in material conformity with the accounting standards and faithfully represent the client firm's underlying financial position and performance (Statement on Auditing Standards [SAS] No. 1, American Institute of Certified Public Accountants, Auditing Standards Board [AICPA] 1972 AU Section 110). Auditors have responsibilities to plan and perform the audit to obtain reasonable assurance that the financial statements are free of material misstatements

[1] There are also agency costs of debt relationships and there are means (e.g. debt covenants) to mitigate these costs (Jensen and Meckling 1976; Watts and Zimmerman 1978). However, agency costs of debt are of less direct relevance to my thesis.

caused by errors or fraud (SAS No. 1, AICPA 1972 AU Section 110 and AU Section 230). In addition to providing assurance, auditors may influence the content of financial reports through the auditor–client negotiation process. If auditors have greater expertise relevant to their clients' operations, they may negotiate more effectively with management regarding the application of the accounting standards and estimates affecting accruals, thereby improving the usefulness of audited financial statements for decision making (Gibbins et al. 2001; Kwon 1996). As investors and their information agents (analysts) rely on the information in the financial statements to frame their understanding of the firm's current position and performance, their predictions of future performance should be more accurate for clients of higher-quality auditors.

The quality of the audit services provided to a client is not observable directly; however, prior studies have identified observable audit firm (or partner) attributes argued to indicate the supply of superior quality audit services. These proxy measures include audit firm size (Francis and Yu 2009; Davidson and Neu 1993), audit firm and/or partner tenure (Chen et al. 2008; Carey and Simnett 2006) and the audit firm's fee dependence on a client (Ball et al. 2012; Lim and Tan 2008). My focus, however, is on audit firm industry specialisation, which I conceive as the extent to which auditors concentrate their business in particular industries, and which thus improves their domain-specific expertise and quality of services rendered (Bedard and Chi 1993). The domain-specific expertise of an industry specialist auditor may arise from more valuable knowledge spillover across clients with fundamentally similar operations (Simunic and Stein 1987; Gramling and Stone 2001). Further, economies of scale in the audit production function may motivate audit firms to specialise by industry to obtain greater benefits from investment in industry-specific technology, training and methodology. Large auditing firms have recognised the importance of specialising in industries and have shown a tendency to concentrate their engagements in particular industries (General Accounting Office 2003b; Gramling and Stone 2001; Hogan and Jeter 1999). Importantly, there is evidence that industry specialist auditors use their superior expertise throughout the audit processes to improve specific audit outcomes. In particular, they are more likely to resist clients' pressure over accrual estimates and policy choices and negotiate with clients to reduce any potential manipulations (Gibbins et al. 2001, Gibbins et al. 2003; Kwon 1996). Prior studies also demonstrate that clients of an industry specialist auditor report lower discretionary accruals (Becker et al. 1998; Krishnan 2003; Balsam et al. 2003), increased disclosure quality (Dunn and Mayhew 2004), lesser likelihood of benchmark-beating earnings management behaviour (Payne 2008) and more timely market responsiveness to earnings surprises (Teoh and Wong 1993; Knechel et al. 2007).

Despite the significant coverage of the effects of auditor industry specialisation on variables indirectly associated with the stated objectives of financial reporting, whether industry specialisation improves the reliability (quality) of financial reports for users' predictions of future performance, the primary criteria identified in the Conceptual Framework, has received less attention. Further, the extant research addressing this issue is compromised by methodological issues. My thesis addresses this gap in our understanding by robustly examining whether the clients of industry specialist audit firms are associated with more accurate analysts' earnings forecasts, and the extent to which there is supporting evidence consistent with the existence of a causal relationship. Below, I briefly describe the key recent literature and identify the methodological weaknesses present in these studies that I propose to address in my thesis.

Some recent empirical research examines the relationship between indicators of high-quality audit services and the predictability of earnings reflected in analysts' short-horizon (end-of-year) forecasts. Behn et al. (2008; hereafter 'BCK') argue that superior audit quality improves the quality of accruals and the reliability of current earnings for predicting future performance. They show that the accuracy of analysts' prediction of earnings increases for clients of Big N audit firms, and that auditor industry specialisation reduces forecast errors, but only for clients of non-Big N auditors. However, Lawrence et al. (2011) find no evidence that clients of Big N audit firms are associated with analyst short-horizon forecast accuracy after matching these firms to non-Big N audit firm clients with similar characteristics. Notwithstanding the above argument and evidence, Payne (2008) conjectures that high-quality audit providers are more likely to constrain management's manipulation of accruals that could be used to 'meet or just beat' the analysts' current consensus forecast (Bannister and Newman 1996; Burgstahler and Eames 2006), which in turn increases realised forecast errors (decreasing forecast accuracy). Payne (2008) shows that auditor industry specialisation (as a proxy for audit quality) decreases the likelihood of 'meeting or just beating' earnings manipulation behaviour [2] and decreases the accuracy of analysts' forecasts of earnings.

[2] Using a matched sample approach, Minutti-Meza (2013) shows that market-share measures of industry specialisation are not associated with any of a broad range of earnings quality proxies, including the likelihood that a client firm 'meets or just beats' the analysts' consensus forecast.

I argue that the above research design, which employs forecasts issued immediately prior to the release of earnings ('short-horizon forecasts'), can only weakly test the relationship between indicators of audit quality, such as auditor industry specialisation, and the predictability of earnings. This is because the accuracy of short-horizon analysts' forecasts is a function of a broad information set, a relatively small proportion of which comprises historic financial reporting data, and is subject to potentially competing effects of audit quality. I aim to explain the conflicting results in the literature and to show that, by employing an alternative research design focusing on the accuracy of forecasts issued immediately after the release of prior period earnings ('long-horizon forecasts'), a robust positive relationship between analyst forecast accuracy and auditor industry specialisation exists. My first research question is:

RQ 1: Is audit firm industry specialisation associated with analyst forecast accuracy?

I contend that the directional effect of auditor industry specialisation on short-horizon forecast accuracy is unclear, but that this indicator of audit quality should improve analyst long-horizon forecast accuracy because these forecasts are less likely to represent a benchmark for the manipulation of earnings at the end of the year. Further, the accuracy of long-horizon forecasts should be more sensitive to information present in prior period financial reports, and thus variations in the quality of the audited financial reports, because this information represents a relatively large proportion of the total information set available at the time of the forecast.

While tests derived from the first research question may produce evidence consistent with auditor industry specialisation causally affecting analyst long-horizon forecast accuracy, I plan to bring more persuasive evidence to bear by investigating whether the empirical association between auditor industry specialisation and analyst forecast accuracy is stronger in cases in which theory suggests that audit quality should have a greater influence on the predictability of earnings. Therefore, my second research question is:

RQ2: Does the strength of the association between audit firm specialisation and analyst forecast accuracy vary with factors affecting the relative importance of audit quality to the predictability of earnings?

I argue that the importance of audit quality varies cross-sectionally with respect to: (a) the underlying riskiness of client firms' operations and (b) the quality (expertise) of analysts covering the auditor's clients. Client firm's operating risk reflects the uncertainty underlying clients' business and financing activities, the level of which may reduce the precision of information available (e.g. accruals) for predicting firms' future cash flows and earnings (Minton et al. 2002). This, in turn, increases the difficulty facing analysts in trying to understand the implications of accruals for future earnings and the difficulty for auditors charged with the responsibility of verifying clients' accrual estimates and policy choices. The effect of auditor industry specialisation on the predictability of earnings should be greater where client's operating risk is higher, because the impact of audit quality on accrual estimation errors should be greater when the scope for such errors in accrual estimates (and thus earnings) is larger. Moreover, high-quality auditors increase their effort in response to increased client risk, with their additional effort having a greater impact in improving earnings predictability (O'Keefe et al. 1994; Schelleman and Knechel 2010; Caramanis and Lennox 2008). Thus, I predict that the effect of auditor industry specialisation on the accuracy of analysts' prediction of earnings is greater when client firm's operating risk is higher. Second, analyst quality captures analysts' superior access to private information beyond that contained in the financial reports as well as their superior ability to identify and process relevant complex information. If auditor industry specialisation increases the reliability of earnings for predicting future earnings as expected, this effect should be pronounced when there is a (greater) need for high-quality earnings and other financial information. Lower-quality analysts rely more heavily on published financial reports to generate earnings forecasts (Das et al. 1998; Clement 1999); accordingly, they should benefit relatively greatly from improvements in the quality of financial reports (and thus audit quality). Higher-quality analysts may benefit from their private information and ability to process complex information, making them less dependent on published financial reports when making forecasts. Therefore, I predict that the impact of auditor industry specialisation on the accuracy of analysts' prediction of earnings will be greater when the quality of the analysts issuing the forecasts is lower.

Below, I briefly describe the research design employed in my thesis. To examine the extent to which audit firms' concentration of their business in particular industries affect the accuracy of analysts' prediction of earnings, I regress the absolute value of short- or long-horizon forecast errors relating to client earnings (inverse measures of forecast accuracy) on measures of audit firm portfolio-share industry specialisation and a vector of control variables (base model). I use audit firm portfolio-share industry specialisation as my main proxy for audit quality because it captures auditors' expertise resulting from audit firms' strategic concentration of their services within industries and it is most closely related to the economic concept of specialisation (Simunic and Stein 1987; Krishnan 2001). Conversely, market-share-based measures of specialisation do not measure specialisation directly, and they are potentially contaminated by undesirable effects associated with market dominance.[3] In examining the extent to which the relationship between audit quality and forecast accuracy varies with client firm's operating risk, I modify and expand the base regression models to include a commonly used proxy for firm's operating risk; that is, cash flow volatility (Minton et al. 2002; Allayannis et al. 2005) and its interaction with audit firm industry specialisation.[4] To explore whether the relationship between audit firm industry specialisation and forecast accuracy varies with the quality of the analysts covering the client firm, I modify the base model and add the main and interaction effects between audit firm industry specialisation, analyst quality and forecast accuracy. In addition to this test, I compute the difference in forecast errors between the 'worst' and 'best' quality analysts, and regress the difference against audit firm industry specialisation and controls. My analyst quality measures are drawn from prior literature (Clement 1999; Chan et al. 2004; Drake and Myers 2011) and include analysts' general experience, firm-specific experience, brokerage size and 'All-Star' status. In addition to these single-attribute analyst quality proxies, I develop two composite measures to capture various aspects of analyst quality. The first composite score is a function of the four above-mentioned analyst quality proxies, while the second includes proxies that represent analysts' personal qualities

---

[3] I test the market-share measures of auditor industry specialisation in robustness tests.
[4] I use the innate accrual quality measure, developed by Francis et al. (2005a) to proxy client firm's operating risk in sensitivity analyses.

(general or firm-specific experience and 'All-Star' status). All of my tests are estimated using both single-stage regression and two-stage regressions, designed to account for the likely endogenous selection of specialist auditors.

I now briefly describe my sample and summarize my main findings. I study U.S. firms for the period 1989 to 2010 and focus on clients of Big N firms to control for any potential factors that might affect my analysis arising from differences induced by audit firm size. In accordance with the existing literature, I find no evidence that auditor industry specialisation consistently improves or impairs analyst short-horizon forecast accuracy, and I show that subtle changes in model specification substantially influence the direction and significance of the measured relationship between these variables. Conversely, I find that long-horizon earnings forecasts for clients of industry specialist auditors are significantly more accurate than are those for other firms, and that these results are robust to numerous model specifications and modelling choices. These findings are consistent with my argument that audit quality has conflicting effects on short-horizon forecast accuracy and that long-horizon forecast accuracy is a theoretically superior meter of the impact of audit quality on the usefulness of financial reports for the purpose for which they are prepared.

Further, I find that the relationship between auditor industry specialisation and analyst forecast accuracy varies cross-sectionally with client firm's operating risk and analyst quality. This variation is consistent with theory regarding the circumstances in which audit quality *should* make a greater difference to forecast accuracy. I show that the relationship between auditor industry specialisation and analysts' absolute forecast errors is stronger when client firm's operating risk is higher, consistent with a greater return to audit quality in situations in which the auditor's task complexity is higher. I further demonstrate that auditor industry specialisation more greatly affects forecast accuracy when analyst quality is lower. This supports my conjecture that the positive effect of audit quality on the usefulness of financial reports for predicting future performance should be greater for lower-quality analysts, who are expected to rely more heavily on published information such as firms' audited financial reports when making forecasts. These findings are broadly robust to a variety of modelling choices, although the results for the 'market-share' type specialisation proxies are less consistent compared to those for my main tests. Overall, my tests of RQ2 (*Does the strength of the association between audit firm specialisation and analyst forecast accuracy vary with factors affecting the relative importance of audit quality to the predictability of earnings?*) provide evidence consistent with the existence of a causal relationship between audit quality and analysts' earnings forecast accuracy.

1.2 Contributions

My thesis has both practical and theoretical contributions. First, I identify the conditions under which published earnings more closely fulfil the forward-looking and 'user coverage' objectives of financial reporting, which is of clear interest to regulators and investors who rely on these reports. In particular, I present evidence that the extent to which an audit firm concentrates its business in an industry is positively associated with the accuracy of analysts' prediction of earnings, consistent with auditor industry specialisation improving the quality of published general-purpose financial reports for predicting future performance (as per SFAC No. 8, FASB 2010 OB2, OB3 and BC1.31). I further show that this improvement in usefulness of financial reports for predicting future performance concentrates among those financial statement users that rely relatively heavily on published financial reports in making decisions (i.e. analysts of lower quality), indicating that financial reports may be useful for a greater number of financially competent users.

Second, I reconcile and explain the inconsistent prior findings regarding the relationship between auditor industry specialisation and analyst forecast accuracy. While other recent papers attempt to provide greater theoretical background to, and explanation for, BCK's hypothesised relation between auditor industry specialisation (audit quality) and forecast accuracy, they are silent regarding the contradictory predictions and findings in Payne's study (Choi and Kwon 2008; He et al. 2011; Lawrence et al. 2011). I show that the directional relationship between auditor industry specialisation and short-horizon analysts' forecast errors is sensitive to model specification, and in particular to the choice of deflator for forecast accuracy and whether the endogenous selection of auditor is controlled. More importantly, I argue and show that the accuracy of analysts' long-horizon forecasts is a superior measure of the impact of audit quality on the usefulness of financial reports. I attribute the greater consistency in the long-horizon results to two factors: (1) long-horizon forecasts are more sensitive to the quality of published accounting data, and (2) managers have less incentive to manipulate future earnings to 'meet or just beat' long-horizon consensus forecasts because these forecasts are likely to be redundant at the end of the reporting year. This is of clear relevance to future research that seeks to investigate the impact of factors posited to affect the quality of published financial reports.

Third, my study contributes more broadly to the auditing literature by enhancing our understanding of the economic role of audit quality. While there is much evidence about the impact of audit quality (proxied by various observable auditor attributes) on management's reporting choices and stock price response (Teoh and Wong 1993; Becker et al. 1998; Francis et al. 1999; Balsam et al. 2003; Knechel at al. 2007; Francis and Yu 2009; Reichelt and Wang 2010), the literature regarding the relationship between audit quality, proxied in my thesis by audit firm industry specialisation, and analysts' decision making is relatively scarce. Audited financial information is a key input in the forecasting process of analysts. My thesis presents evidence that industry specialist audit firms do improve the accuracy of analysts' forecasts, with the effect being more pronounced when firm's operating risk is relatively high and the quality of analysts is relatively low. These results produce strong evidence consistent with the existence of a causal relationship between audit quality and analysts' forecasting performance.

Finally, my study contributes to the financial analyst forecasting and earnings quality literature. Sell-side analysts are sophisticated intermediaries, whose earnings forecasts appear to be strong proxies for market expectations of future earnings (Kothari 2001; Affleck-Graves et al. 2002; Crabtree and Maher 2005; Dichev and Tang 2009). My findings should be useful to investors and creditors who make decisions upon analysts' forecasts, as my results suggest that analysts' forecasts are more accurate for client firms that are audited by an industry specialist audit provider. Further, the firm-year-level analyst quality metrics developed in my thesis may be of use in testing the relative impact of other earnings quality proxies. In particular, the difference in forecast accuracy between the 'worst' and 'best' quality analysts following a given firm may be usefully applied in assessing the impact of real and accrual-based earnings manipulation and disclosure quality.

1.3 Outline of the Thesis

The remainder of my thesis is organised as follows. Chapter 2 reviews the literature of general and specific relevance to my research questions. This chapter develops and describes a theoretical framework, based on that of Knechel et al. (2013), which articulates the means by which the supply of higher-quality audit services may lead to superior financial reporting quality. While my literature review emphasises the connection between audit firm industry specialisation and measures of audit or financial reporting quality, the relevant literature concerning audit quality and its impacts more broadly is also analysed. Building on the extant literature, I develop five hypotheses in Chapter 3. The first two hypotheses predict an association between audit firm industry specialisation and analyst short- and long-horizon forecast accuracy, while the remaining three hypotheses propose that the relationship between industry specialisation and analyst forecast accuracy varies cross-sectionally with the underlying riskiness of the client firms' operations and the quality of the analysts covering the client firm. Chapter 4 introduces and explains the empirical measures chosen as my test variables: audit firm industry specialisation and analyst quality. I identify and describe several measures of audit firm industry specialisation and argue that audit firm portfolio-share industry specialisation is the most direct measure of the underlying phenomena that I wish to study. I further describe several measures of analyst quality proxies that are commonly used in the prior literature and identify analysts' general and firm-specific experience, brokerage size and *II* 'All-Star' status as the measures most relevant to my study. Chapter 5 describes the research design, starting with an overview of the general form of the regression models, followed by a description of the variable measurement and estimation methods, and concluding with a detailed description of the full regression models employed to test my hypotheses. A description of the samples used to test my hypotheses is provided in Chapter 6. I then report and analyse the results for the tests of the first two hypotheses in Chapter 7 and the remaining three hypotheses in Chapter 8. Chapter 9 discusses the conclusions, implications of my study, limitations and future research opportunities.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

My thesis examines the relation between audit firm industry specialisation and client financial reporting quality, as measured by the accuracy of securities analysts' forecasts of client earnings. In this chapter, I review and analyse the findings of the relevant literature to inform the later development of specific hypotheses relating to my research questions. Central to my research questions is the issue of whether industry specialist audit firms provide superior quality audit services to their clients. Consequently, my literature review encompasses studies of the broader nature and impacts of audit quality in addition to those focusing on industry specialisation.

This chapter is organised as follows. The underlying audit quality constructs employed in the existing literature are described and discussed in Section 2.2. In Section 2.3, I develop a theoretical framework modelled on that of Knechel et al. (2013) to identify the mechanisms through which the supply of higher-quality audit services may lead to higher-quality audit outcomes. Specific applications of this framework are elaborated with supporting empirical evidence in Sections 2.4 – 2.7. Section 2.4 identifies and analyses the elements of audit inputs that are argued to improve auditors' judgment and negotiating power and consequently improve financial reporting quality. The extant empirical evidence regarding these audit inputs and auditors' behaviour during audit processes is reviewed in Section 2.5. In Section 2.6, I examine auditor incentives and their influence on the above relationship. Evidence on how audit inputs and processes affect audit outcomes is discussed in Section 2.7. Section 2.8 concludes the chapter. Chapter 3 then applies these general findings to the development of specific hypotheses concerning the relationship between audit firm industry specialisation and analyst forecast accuracy.

2.2 Audit Quality Constructs

Academics and regulators have proposed several definitions of audit quality over the past three decades. Although 'audit quality' is the subject of an extensive scholarly literature, the underlying construct considered and the empirical proxies employed differ significantly across the literature. This lack of scholarly consensus is reflected in regulatory pronouncements recognising the difficulty in defining audit quality and conceding that the regulation of auditor performance against any assumed quality definition is problematic (Financial Reporting Council [FRC] 2006, 16; International Organisation of Securities Commissions [IOSCO] 2009, 3).

Two related but distinct constructs underpin scholarly audit quality studies, with each reflecting a different aspect of the services rendered by an auditor to their client. The first construct conceives audit quality singularly as a function of the likelihood that material misstatements and errors in the client's financial statements will be detected and reported (DeAngelo 1981). However, the detection and reporting of errors and irregularities does not necessarily imply that the objectives of financial reporting are being met.[5] My thesis employs a broader audit quality construct that emphasises the auditor's impact on the overall quality of the client's financial reports, where reporting quality reflects the attributes of information that are considered

[5] Recall that the objectives of financial reporting are to provide financial information that is useful for investors, lenders and creditors' allocations of scarce resources, assessing firms' future cash flows and future performance (SFAC No. 8, FASB 2010 OB1-3, and BC1.31).

desirable by regulators, scholars or financial statement users (Palmrose 1988; BCK 2008; DeFond and Zhang 2014). Studies invoking this 'financial reporting quality' construct typically consider the auditor's role in providing assurance on, and improving the relevance and reliability of, the audited financial statements (DeFond and Zhang 2014). This effect may derive from the auditor's influence on management's accounting estimates and policy choices, in addition to simply detecting and reporting errors (Francis and Yu 2009).

2.3 Auditor Attributes, the Audit Processes and Financial Reporting Quality

The broader audit quality construct is operationalised by observing audit outcomes: financial reporting data and various financial reporting quality summary measures are commonly used. Since the impact of an audit/auditor on reporting quality is rarely observable directly, prior studies focus on the relationship between observable audit firm (or partner) attributes[6] and proxies for financial reporting quality (e.g. analyst forecast accuracy, accounting restatements, discretionary accruals, and disclosure quality). In this section, I first identify and briefly describe various financial reporting outcomes argued to indicate superior quality audit services, and then introduce and describe a theoretical framework identifying the mechanisms through which the observed audit/auditor attributes may affect financial reporting outcomes. Subsequent sections of this chapter describe in detail the theoretical and empirical literature concerning the audit inputs (Section 2.4), processes (Section 2.5),

---

[6] Common proxies are audit firm size (Davidson and Neu 1993; Francis and Yu 2009), audit firm industry specialisation (Krishnan 20003; Payne 2008), auditor tenure (Chen et al. 2005; Carey and Simnett 2006) and audit fees (Ball et al. 2012; Lim and Tan 2008).

moderating factors (Section 2.6) and outputs (2.7) identified in the Audit Quality Framework.

2.3.1 Audit Outcomes—Financial Reporting Quality

My thesis maintains that audit quality can be inferred by observing financial reporting quality. In this sub-section, I briefly introduce various measures of financial reporting quality, to contextualise the following discussion of the Audit Quality Framework. In a later section (2.7), I return to examine the relation between audit inputs and processes and financial reporting quality more deeply.

Broadly, I assume that the quality of financial reporting reflects its consistency with the objectives of financial reporting as prescribed in the Conceptual Framework, which states: 'The objective of general purpose financial reporting is to provide financial information about the reporting entity that is useful to existing and potential investors, lenders, and other creditors in making decisions about providing resources to the entity' (SFAC No. 8, FASB 2010 OB1 and OB2). Central to these decisions is the need to assess the prospects for future net cash inflows to an entity (SFAC No. 8, FASB 2010 OB3). The Framework also acknowledges that the provision of information about a firm's financial performance as measured by earnings and its components is the primary focus of financial reporting, as earnings information is considered superior to current cash flow information for the purpose of predicting future cash flows (SFAC No. 8, FASB 2010 BC1.31).

The empirical reporting quality literature assesses the extent to which these objectives are achieved by measuring financial statement (audit) outcomes, which comprise: (a) measures of the properties of the financial reports and related disclosures (e.g. disclosure quality and accrual manipulations) and (b) measures of financial statement users' reaction to financial reports (e.g. stock price response to accounting information and analyst forecast accuracy).

Measures reflecting the properties of the financial reports and accompanying disclosures include disclosure quality and accrual manipulations. Disclosure quality encompasses the quantity and decision usefulness of client disclosures additional to the data presented on the face of the financial statements. Such disclosure includes both financial (e.g. management forecasts of earnings or sales) and non-financial information (e.g. information on ownership structure, new products and business plans). Discretionary accruals, estimated as the proportion of a client's total accruals that cannot be explained by the client's economic fundamentals, are argued to capture management's discretionary reporting decisions. Greater absolute discretionary accruals are argued to reflect poorer financial reporting quality (Schipper and Vincent 2003). Payne (2008) implicitly views analyst forecast accuracy as a reporting outcome, arguing that the manipulation of reported earnings allows clients to 'meet or just beat' consensus forecasts. From this perspective, forecasts that are more accurate are argued to reflect greater earnings manipulations and thus poorer audit quality.

Measures reflecting investor reaction to financial reports include aggregate stock price response and analyst forecast accuracy. Market responsiveness to earnings surprises reflects the market perception of clients' earnings quality, and is measured by the speed and bias in the price reaction to earnings news (Teo and Wang 1993;

Kothari 2001). Other studies focus on the impact of audit quality on client cost of capital (Khurana and Raman 2004). Unlike Payne (2008), authors such as BCK (2008) and Choi and Kwon (2008) view greater analyst forecast accuracy in a positive light, arguing that accuracy reflects how effectively sophisticated users of financial reports are able to use reported financial information to predict future realisations of earnings. Thus, in this sense, analyst forecast accuracy is viewed as a *response* to the quality of financial reporting (BCK 2008; Choi and Kwon 2008), and may directly assess the extent to which the financial reporting objectives are achieved.

2.3.2 Theoretical Audit Quality Framework

There have been several recent attempts to develop a framework within which to understand audit quality and audit quality research. The U.K. Financial Reporting Council (FRC 2008) proposes an Audit Quality Framework designed to complement existing regulations and to promote five key drivers of audit quality: (1) the organisational culture within an audit firm, (2) the skills and personal qualities of audit partners and staff, (3) the effectiveness of the audit process, (4) the reliability and usefulness of audit reporting and (5) factors outside the control of auditors.

More recently, scholars including Francis (2011) and Knechel et al. (2013) propose frameworks for understanding audit quality and related research. Francis's (2011, 125) framework is intended to help scholars, practitioners, regulators and policymakers to understand better the various drivers of audit quality. Francis (2011) argues that audit quality is affected by six units of analysis, beginning with a granular view of the audit inputs, to a broader view of the economic outcomes of an audit. These units of analysis include audit inputs, audit process, accounting firm (i.e. engagement team), audit industry and markets (i.e. industry structure), institutions (i.e. State Boards of Accountancy, the AICPA, FASB, SEC and PCAOB and legal system) and economic consequences of audit outcomes. Knechel et al. (2013) pursue a similar objective to Francis (2011) and propose a framework based on a large body of auditing research to conceptualise audit quality as arising within a system comprising inputs, process, context and outcomes.[7] While these two frameworks are alike, the Knechel et al.'s (2013) framework (hereafter 'KKPSV Framework') allows a clearer focus on the various audit inputs and processes and the links therein, and is thus employed as the base of the framework used in my thesis.

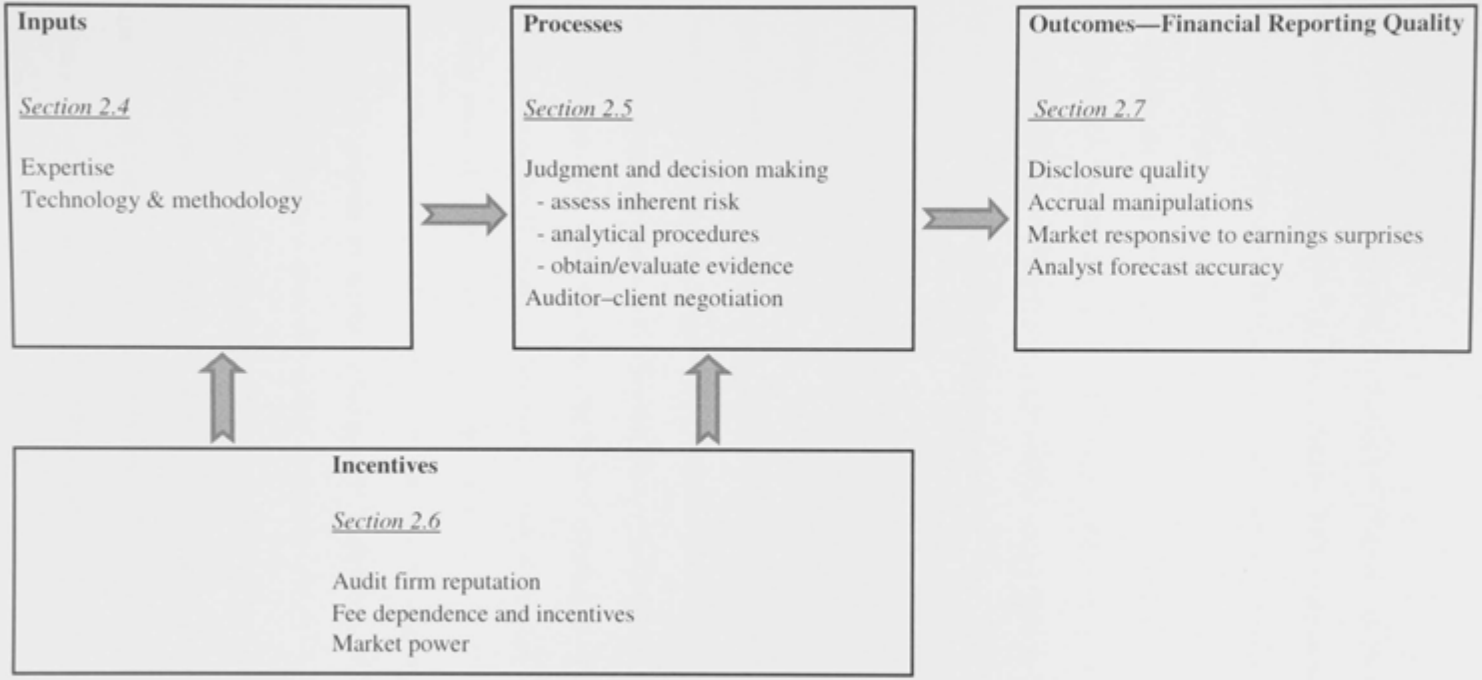
While the frameworks described above are useful for understanding and classifying a diverse audit quality literature, they are less well suited to explaining the link between observable auditor attributes and audit outcomes. Consequently, I modify the KKPSV Framework to accommodate an examination of the link between observable auditor attributes, such as industry specialisation, and financial reporting quality. The framework within which I explain the theory underpinning my research question is illustrated in Figure 2.1.[8]

The logic underpinning my audit quality framework reflects the following maintained assumptions and definitions. First, audit quality is innately related to the quality of the resulting financial reports, which in turn reflects the extent to which these reports satisfy the reporting objectives identified in the regulatory conceptual frameworks. Second, the auditor's impact on reporting quality can be directly

[7] The original KKPSV framework is presented in Appendix A.
[8] The Framework is described and elaborated with supporting empirical evidence in Sections 2.4–2.7.

assessed by observing audit outcomes, which are a function of imperfectly observed audit inputs and processes. Audit inputs ('Inputs' in Figure 2.1) comprise the human and technological resources dedicated to an audit, and include the expertise of audit personnel and the effectiveness of technological and methodological support resources. Audit processes ('Processes' in Figure 2.1) consist of a series of activities leading to the issuance of an audit opinion, and include analytical review procedures, control tests, substantive tests, negotiation with clients and quality control procedures. The application of professional judgment and decision making is critical in each audit process. In particular, audit personnel exercise judgment in performing analytical procedures, assessing clients' risks and obtaining and evaluating evidence. After assessing controls and performing substantive audit tests, auditors negotiate with clients. The auditor's ability to influence clients' reporting decisions regarding the accounting estimates and choices presented in their financial reports may affect financial reporting outcomes, and thus audit quality. However, while an auditor may possess great expertise and have access to superior support resources and thus be capable of exercising high-quality judgment and negotiation in audit processes, personal incentives ('Moderators' in Figure 2.1)—such as reputational concerns, fee dependence and market dominance—may affect the extent to which inputs and professional judgment are actually applied to audit processes. This, in turn, has the potential to affect the quality of financial reports. Each of the above key inputs and processes and their relationships with each other and with audit outcomes are examined in detail in the following sections.

Figure 2.1 Audit Quality Framework



2.4 Inputs

Audit inputs comprise the human and other resources dedicated to an audit, and include audit staff expertise and audit support resources. While experimental studies typically examine the association between inputs and the quality of audit processes, archival research largely focuses on the influence of the observable attributes of inputs on proxies for financial reporting quality. As the interrelationship between inputs, processes and outcomes is complex, I first describe the audit inputs identified in the literature in this section and highlight that industry specialist auditors possess these inputs. Later sections detail the interrelationships between inputs, processes and outcomes.

2.4.1 Expertise

A key input identified in the audit literature is the collective expertise of the audit staff employed. Expertise is generically defined as a high level of knowledge or skill in a particular field (*Oxford Dictionary of English* 2010, 616; *Cambridge Advanced Learner's Dictionary* 2008, 492). Scholars have defined expertise as consisting of 'knowledge about a particular domain, understanding of domain problems, and skill about solving some of these problems' (Hayes-Roth et al. 1983, 400), where knowledge is defined as 'acquired information that can be activated in a timely fashion in order to generate an appropriative response' (Charness and Schultetus 1999, 6) and skill refers to a particular ability developed through training and experience to do something well (*Cambridge Advanced Learner's Dictionary* 2008, 1347). In auditing research, this ability is frequently contextualised as problem-solving ability (Bedard and Chi 1993; Bonner and Lewis 1990). Therefore, my thesis defines auditor expertise as an auditor's knowledge about auditing problems and skill in solving these problems. Knowledge and skills are considered to be developed through training and practice.

Since expertise is not directly observable, the literature employs proxy measures for this purpose. In early behavioural research, expertise is measured as an individual's years of experience in a discipline, with the assumption that knowledge and skills are gained through years of practice. This perspective has been adopted in clinical psychology (Oskamp 1965), physics (Chi et al. 1982) and auditing (Ashton 1991; Frederick and Libby 1986; Hamilton and Wright 1982).

However, this experience-based proxy for expertise has been criticised in recent studies, which contend that auditors' expertise is domain-specific. Knowledge and skills may not be simply accumulated with years of experience; the nature of that experience matters. For example, Bedard and Chi (1993) propose that auditors need to possess general auditing knowledge as well as knowledge about accounting principles and client-industry specifics (domain-specific knowledge). Auditing a manufacturing company requires a different set of skills and knowledge than that applied to the audit of an insurance company, because the transaction structure, tax rules and accounting systems are substantially different. Relevant domain-specific knowledge may be gained through serving many clients in that industry or auditing clients in a particular industry for many years, rather than simply performing audit work for many years. A similar argument is proposed by Bonner and Lewis (1990) regarding problem-solving abilities; they further point out that experience together with ability develops knowledge, and that knowledge combined with ability improves performance.

Accepting that the effect of experience on expertise is contingent, recent experimental auditing researchers (Solomon et al. 1999; Wright and Wright 1997; Taylor 2000; Low 2004; Moroney 2007)[9] and archival researchers (Davidson and Neu 1993; Balsam et al. 2003; Krishnan 2005; Reichelt and Wang 2010)[10] use proxies such as audit firm size or industry specialisation—which are argued to indicate greater knowledge and problem-solving skills—when studying the effect of expertise on judgment and financial reporting quality.[11]

Audit firm size and industry specialisation are argued to proxy expertise for a number of reasons. Large audit firms are argued to recruit greater numbers of audit personnel, and personnel with superior expertise. DeAngelo (1981) argues that Big N firms have greater industry-specific knowledge and possess more expertise in preparing SEC documents relative to small firms. In addition, large audit firms have a greater incentive and ability to promote in-house skills and knowledge transferring among auditors. Francis and Yu (2009) conjecture that these expertise accumulating and sharing processes are more profound in larger audit offices, as such firms/offices have more substantial training programs, standardised audit programs and information technology support. A similar argument is apparent in studies that examine industry specialist audit firms. Specifically, Solomon et al. (1999) claim that industry specialists have superior knowledge of the frequency of financial statement error in industries in which they are specialised relative to other industries. In addition to general auditing knowledge, individual auditors working for industry specialist firms can gain domain-specific knowledge about their client's industry

---

[9] Experimental researchers use industry-related experience to measure industry specialisation.
[10] Archival researchers also employ industry specialisation to proxy expertise and use various industry specialisation measures, which I describe substantively in Chapter 4.
[11] The empirical evidence is elaborated in the following sections.

(Bedard and Chi 1993) and may develop this expertise relatively quickly, as they can access training that is more comprehensive and can learn from colleagues and seniors that have significant industry-specific experience.

The impact of expertise on performance has been studied in many contexts (e.g. chess masters' recall superiority and physics problem solving speed), with a positive relationship generally being documented (Chase and Simon 1973; Simon and Simon 1978). The impact of expertise on auditors' judgment quality, negotiation powers over clients and opinions on financial reports has also been extensively researched in the auditing literature. This is examined in detail in Section 2.5.

2.4.2 Technology and Methodology

Technological and methodological support is another important audit input that potentially affects the quality of the audit services supplied. In the previous section, I discussed why auditor expertise might affect performance. To be an expert, knowledge is essential. However, expertise is not innate; it is acquired through practice, with instruction, training and feedback. Audit firms assist audit staff in the collection and analysis of evidence by devising audit programs, testing procedures and providing internal administrative support (Francis 2011). Thus, an audit firm's comprehensive support system (including technological and methodological support) facilitates the knowledge and skill acquisition process, in turn improving auditors' judgments and quality of decision making throughout the audit processes.

An audit methodology is a particular set of processes and procedures that guide auditors from the preliminary risk assessment phase to the reviewing phase. It is designed to help auditors to cope with uncertainty in a systematic manner (Knechel et al. 2013). Public accounting firms deploy different audit methodologies, and this may affect auditors' judgment (Wilks and Zimbelman 2004). Advanced information technology and audit programs provide essential technical support to auditors and may directly influence the knowledge acquisition process and auditors' judgment, which will ultimately affect audit outcomes (Dowling and Leech 2007; O'Donnell and Schultz 2003; Janvrin et al. 2008).[12]

Large and industry specialist audit firms are argued to provide auditors with greater technological and methodological support. Specifically, because large audit firms may benefit from the economies of scale or scope in the audit process, they may invest heavily in specific technologies and/or methodological support. Prior study suggests that large audit firms (proxied by Big N membership) develop their own methodological framework complying with the accounting rules and auditing standards to guide auditors, who will therefore be more likely to make better judgments in the audit processes and increase audit outcomes (Davidson and Neu 1993). Audit firms of given size are more likely to invest in technologies, physical facilities, personnel and organisation control systems in industries in which they have a greater likelihood of obtaining a return on that investment. For auditors with many clients in an industry, the potential returns to industry-specific investment are greater (Simunic and Stein 1987; Gramling et al. 2001). Therefore, it is reasonable to assume that industry specialist firms provide superior audit support resources, which improve judgment and decision-making processes and ultimately lead to high-quality audit outcomes.

---

[12] The empirical evidence regarding how technology and methodology improve judgment and decision-making quality is reviewed in Section 2.5.

2.5 Processes

Audit processes comprise a series of activities leading to the issuance of an audit opinion. This includes analytical review procedures, control tests, substantive tests, client negotiation and quality control procedures. The application of professional judgment and decision making is critical at each step in the audit processes. Audit services are not homogenous; they are provided to diverse clients with different reporting risks. The idiosyncratic nature of an audit service requires resources to be tailored to each client during each stage of the audit. Therefore, the quality of the auditor's judgment and decision making applied during the various stages of the audit and the quality of negotiation outcomes are argued to affect audit outcomes.

Audit inputs and their importance have been discussed in the previous section. Generally, auditors with greater expertise and better support are likely to apply superior professional judgment in assessing audit evidence, which in turn increases their power over clients in negotiations regarding reporting matters (Wright and Wright 1997; Taylor 2000; Chen et al. 2005; Moroney 2007). This section explains the mechanisms through which superior quality audit inputs are considered to improve the quality of audit processes. The empirical evidence is also reviewed with regard to the relationship between various indicators of audit input quality and the quality of audit processes, with an emphasis on the evidence of the role of industry specialisation in this context.

2.5.1 Judgment and Decision Making

Bonner (2008, 2) describes judgment as 'forming an idea, opinion, or estimate about an object, an event, a state or another type of phenomenon', whereas decisions reflect 'making up one's mind about the issue at hand and taking a course of action'. Crucially, decisions taken reflect judgments. As such, the quality of judgments is vital to the quality of decision outcomes. In auditing research, judgment and decision making quality is vital to superior audit outcomes. This is because auditors apply judgment in making decisions throughout all stages of an audit, and their judgment and decision making quality have an immediate and direct influence on clients' financial statements (Knechel et al. 2013, 13). Empirical research (described below) demonstrates that auditor expertise and comprehensive audit support systems promote auditors' judgment quality in performing analytical procedures, assessing the client's risk and obtaining and evaluating audit evidence.

2.5.1.1 Assessing risk

Auditors' risk assessments are important because they can have a significant effect on the subsequent nature, extent and conduct of an audit (Arens et al. 2013, 233; Knechel et al. 2013). Industry specialists are argued to possess greater industry-specific expertise, which may allow them to identify more accurately client's material misstatements. For example, Low (2004) examines the impact of industry specialisation (as measured by industry-related experience) on auditors' judgment and decision making at the planning stage, and finds that industry specialists can efficiently utilise knowledge of the client's industry and more accurately assess the audit risks associated with a client. Low's results also show that industry specialists are associated with a greater likelihood of modifying the nature of the planned audit procedures, increasing the quality of audit procedure changes and final audit programs. Taylor (2000) presents experimental evidence that banking industry specialists are less conservative in assessing clients' inherent riskiness and are more confident about their assessments than are non-specialists with similar experience.

I have previously argued that large audit firms and industry specialists are likely to employ the methodologies best suited to their industries and that these methodologies may affect their assessments of client firms' inherent risk. Studies have found that the audit methodology applied can affect auditors' assessments of inherent risk, including that attributable to fraud risk. Wilks and Zimbelman (2004) study 95 audit managers from Big 5 accounting firms and find that auditors' fraud risk assessments are more likely to be understated when they use a holistic fraud risk-assessment approach rather than using an approach that separates assessments for management's attitude, opportunities and incentives. Therefore, to the extent that industry specialists use a superior methodology, they may provide risk assessments that are more accurate, given that audit methodology improves the accuracy of risk assessment.

2.5.1.2 Conducting analytical procedures

Analytical procedures, required during the planning and completion phases of an audit, are used to assess the reasonableness of account balances or other data through comparisons and relationship analysis (Arens et al. 2013, 132). Prior research investigates the importance of audit support systems on judgment and decision making quality during analytical procedures. O'Donnell and Schultz (2003) examine the influence of different audit support software on auditor planning-phase analytical procedures, and find that auditors who use business-process-focused software identify more of the seeded risk factors and assess misstatement risk at higher levels than their counterparts who use transaction-cycle-focused software. They argue that the business-process-focused format provides a more effective framework for auditors to acquire knowledge, bringing a subsequent improvement in their decision making quality. This is consistent with prior claims and evidence that decision-making aids embedded in audit support systems (rather than checklists)[13] can enhance audit quality through improving audit efficiency and risk management and promoting compliance with accounting standards and the firm's methodology (Ashton and Willingham 1988, Dowling and Leech 2007).

Other research examines the relationship between auditor expertise and judgment and decision making quality in analytical procedures. By accumulating knowledge and problem-solving abilities, experts become aware of the relations of internal control weaknesses and accounting errors and the relations among accounts, and are more likely to make judgments that are consistent with accounting and auditing theory (Frederick and Libby 1986). Using industry-related experience to differentiate experts and novices, Wright and Wright (1997) conduct an experiment on 34 auditors with significant retailing experience and 38 auditors without such experience. They argue that industry specialists enhance hypothesis generation in the planning phase, which leads to a greater likelihood of industry-specific error detection. This improves the effectiveness and efficiency of subsequent audit testing and ultimately leads to superior decision performance. Wright and Wright's results support these expectations, which are also consistent with early evidence generated by Bedard and Biggs (1991), who report that auditors with greater industry

[13] There are various types of decision aids, including checklists, knowledge-based expert systems, decision support systems. Decision support systems are interactive computer-based software that assist auditors in making decisions.

experience (as a proxy for specialisation) are more likely to identify the errors present in a complex analytical procedure task.

2.5.1.3 Obtaining and evaluating evidence

An important audit process involves the identification and evaluation of relevant client data. Auditors with greater expertise are argued to be better able to link concepts learned, organise them to develop effective problem-solving strategies and apply them to actual tasks (Bonner 2008). Therefore, these auditors are more likely to increase information search efficiency and the effectiveness of judgments in the evidence evaluation phase. For example, Bedard and Mock (1992) use a computerised information display board with more than 200 information items to study audit experts' and novices' information acquisition behaviour, and report that experts obtain a greater amount of relevant information and a lesser amount of irrelevant information compared to novice auditors.

Using auditor's industry-related experience as a proxy for expertise, Moroney (2007) reports that industry specialists outperform non-specialists with regard to the time taken to read case material, search for and read information cues and in the effectiveness of information usage (as measured by consistency with the expert panel's model solutions to the hypothetical cases). These results are stronger for clients in the superannuation industry than for manufacturing clients, which is attributable to the greater client heterogeneity in the latter industry. Moroney and Carey (2011) present experimental evidence that industry-based experience is more important than task-based experience in improving auditor performance (measured by the completeness of experimental audit participants' solutions when compared to those provided by the expert panel).

2.5.2 Auditor–Client Negotiation

Auditors negotiate with clients throughout the audit processes regarding the choice of appropriate accounting policies, estimates and disclosures. Audited financial statements are the product of the auditor–client negotiation process (Antle and Nalebuff 1991). As such, the negotiation process plays an integral role in determining financial reporting quality. However, several factors may influence the extent to which the published financial reports reflect clients' or auditors' preferred position. Prior research (e.g. Gibbins et al. 2001; Gibbins et al. 2003; Brown and Wright 2008) provides comprehensive reviews of the negotiation research in auditing, and a number of factors have been identified as affecting negotiation outcomes. These are external conditions and environmental characteristics (e.g. GAAP, regulatory bodies, and litigation risk), the interpersonal context (e.g. auditor–client relationship, personal and organisational agendas, and expectations), client characteristics (e.g. expertise, risk tolerance, and power) and auditor characteristics (e.g. auditor accounting expertise and negotiation experience).

The audit firm's industry expertise is one of the most influential factors affecting the auditor–client negotiation process and outcomes (Gibbins et al. 2001, Gibbins et al. 2003). To the extent that industry specialist firms have greater expertise relevant to their client's industry, they will be more confident in their propositions and demonstrate a greater ability to resist management's pressure over financial reporting issues and constrain management's discretion in applying accounting principles

(Kwon 1996). In this case, the audited financial statements are more likely to be of high quality and be presented consistent with the auditor's preferred position, rather than that of the client (Gibbins et al. 2001; Chen et al. 2005). Instead of directly assessing auditors' expertise, Chen et al. (2005) examine clients' perception of auditors' expertise, presenting survey evidence that auditors are more likely to succeed in the negotiation process when they are perceived to be industry specialists.[14]

2.6 Moderators/Incentives

Like all human actors, auditors are subject to private incentives that may diminish or enhance the extent to which auditor expertise actually flows through to superior process performance and thus superior audit quality. These incentives, discussed in turn below, include reputational concerns, fee incentives and market dominance incentives.

2.6.1 Audit Firm Reputation

Fombrun et al. (2000, 87) consider corporate reputation the 'cognitive representation of a company's actions and results that crystallises the company's ability to deliver valued outcomes to its stakeholders'. Reputation is gained though constantly providing products or services that are superior to those of other companies and that exceed people's expectations. Prior research argues and shows that a good company's product differentiation strategy is important to maintain reputation

[14] Chen et al. (2005) use clients' CFO or CAO perception of an audit firm's industry specialisation to measure industry specialisation and document evidence supporting their expectations. In their sensitivity tests, they use audit firm industry market share (detailed in Chapter 4) to measure industry specialisation, and generate inconsistent results. They explain that the differences in their results may be attributable to the importance of clients' perception of an auditor's industry specialisation in the complex negotiation process.

among its customers, and this will have a positive impact on its pricing power (Jarmon 2009).

Audit firms are argued to invest in brand name capital to signal quality. The established brand name acts as a bond guarantor of auditors' performance so the audit firm can earn quasi-rents on its reputation (DeAngelo 1981). Thus, firms with established reputation are more likely to deliver higher-quality services, as the private cost of supplying low audit quality services is greater than the associated short-term benefits.

Large audit firms are particularly concerned about their reputation capital. These firms are internationally renowned and subject to significant public attention. Further, the greater wealth of large audit firms provides a relatively strong incentive to maintain service quality considering the low importance of one single client weighed against the risk that the audit firm could lose entire clients if caught misreporting for one client (DeAngelo 1981; Francis and Wilson 1988). Several studies that report significant fee premia to Big N audit firms attribute part of this premia to a reputational effect (Francis 1984; Francis and Simon 1987; Bandyopadhyay and Kao 2001). Applying this reputation hypothesis, empirical research documents a positive association between Big N audit firms and financial reporting quality, such as disclosure quality (Davidson and Neu 1993) and earnings quality (Teoh and Wong 1993; Francis and Yu 2009).

Similarly, to the extent that specialisation allows audit firms to provide a differentiated product, industry specialist firms may experience stronger incentives to protect their reputation capital, and thus economic rents, by providing higher-quality audit services (Craswell et al. 1995; Abbott et al. 2005). Industry specialist firms may develop their reputation by investing in acquiring industry-specific expertise over time, which may translate into fee premia. Craswell et al. (1995) show that audit fee premia are higher for industry specialists than non-specialists within a sample of Australian Big N firms. Similar results are documented by DeFond et al. (2000) in a Hong Kong setting.

2.6.2 Fee Dependence and Incentives

Fee dependence refers to the extent to which a client (consciously or otherwise) exerts power over an audit firm (or audit partner) due to the relative importance of that client's fees in the auditor's (partner's) portfolio. If a client's fees represent a large proportion of an auditor's total fee income, an incentive may exist to lower audit quality to avoid losing the client. While the empirical evidence regarding the impact of fees on audit quality is mixed, there is evidence that fee dependence may impair auditor judgment and decrease financial reporting quality (Houston 1999; Chen et al. 2010; Frankel et al. 2002). Houston (1999) experimentally examines the influence of fee pressure and client risk on audit seniors' judgment and decisions in time budgeting, showing that these decisions are less responsive to increased client risk when auditors are subject to fee pressure. Chang and Hwang (2003) conduct a similar experimental study and show that auditors may tolerate clients' aggressive reporting practices (i.e. auditors tolerate clients' estimation of bad debt expenses and accept clients' footnote disclosure rather than proposing an accounting entry that decreases clients' earnings) when clients' retention incentives are high and clients' business risk concerns are low. Chen et al. (2010) demonstrate that auditors' propensity to issue a modified audit opinion is lower for clients whose fees are relatively important to audit partners when investor protection is relatively weak.

Other research finds that fee incentives or client importance does not compromise auditors' independence or reduce audit quality. For example, Big N audit firms are found to be more conservative with respect to clients' discretion on accruals (Reynolds and Francis (2001) and insurance clients' tendency to under-reserve (Gaver and Paterson 2007) when these clients are more important (and thus pay greater audit fees) to the local audit offices. Chen et al. (2010) report that the negative relationship between the probability of an individual auditor's issuance of a modified opinion and client importance (greater audit fees) becomes insignificant during periods of improvement in China's legal and regulatory environment. This evidence is consistent with the argument that auditors (particularly those with large clients) will trade off economic dependence for reputation protection and litigation avoidance (Reynolds and Francis 2001; Chen et al. 2010; Stice 1991).

The provision of non-audit services may further strengthen auditors' economic bond to their clients. Frankel et al. (2002) report a positive association between non-audit fees and firms' absolute discretionary accruals (a measure of earnings management). However, other studies refute this association (Francis and Ke 2003; Ashbaugh et al. 2003; Larcker and Richardson 2004), with some even providing an opposite finding to Frankel's (e.g. Reynolds and Francis 2001). Moreover, DeFond et al. (2002) find no evidence that non-audit services threaten audit outcomes when outcomes are surrogated by auditors' propensity to issue going-concern opinions. These later studies conclude that reputational incentives and litigation concerns dominate fee retention incentives.

As discussed above, industry specialists are strongly motivated to protect their reputational capital. It is thus reasonable to contend that their reputational incentives outweigh any fee incentives, maintaining their judgment and decision quality.

2.6.3 Market Power

Economic theory suggests that monopoly power exists where demand and marginal revenue are divorced (Taylor and Frost 2009, Chapter 8). There have been continuing concerns regarding the dominance of the audit market by the Big 4 firms, and the potential negative impact of Big 4 market power on the quality of audit outcomes in the U.S., Great Britain and the European Union (House of Lords 2010a, 2010b, 2010c; European Commission 2010; General Accounting Office 2003a). Specifically, the U.S. General Accounting Office (2003a, preface) states that 'domestically and globally, there are only a few large firms capable of auditing large public companies, which raises potential choice, price, quality and concentration risk concerns'. Although the General Accounting Office finds no evidence of impaired audit quality, it raises concerns regarding the potential negative effects from market concentration and the market power of the major audit firms.

Recent academic research has also addressed this issue of market concentration by providing both theoretical and empirical evidence. One perspective is that the perceived lack of competition in the audit market may reduce the incentives for dominant auditors to conduct high-quality audits; however, others argue that the observed Big 4 dominance *reflects* a demand for high-quality audits. Numan and Willekens (2012) examine the association between audit quality, competitive pressure and industry specialisation, and find that audit quality has been negatively affected by pressure from close competitors (a measure based on auditors' industry market share), not by industry specialisation *per se*. Francis et al. (2012) report that audit quality is decreased in countries with greater within-Big 4 market concentration, but is increased when the Big 4 have a larger market share as a group. In summary, based on the above regulatory concerns and competing evidence, the use of market concentration as an indicator of audit quality is problematic. My thesis measures audit quality in an alternative way: by capturing auditors' expertise within their own portfolios, rather than relying on auditors' market dominance. I detail this measure in Chapter 4.

2.7 Audit Outcomes—Financial Reporting Quality

As noted in Section 2.3.1, I assume that audit quality can be observed or inferred from financial reporting quality indicators. Financial reporting quality reflects the usefulness of financial reports for the prediction of future earnings (and thus long-run cash flows) (SFAC No. 8, FASB 2010 BC1.31). This aspect of financial reporting quality can be measured by observing (a) the properties of financial reports and accompanying disclosures, such as disclosure quality and accrual manipulations, and (b) the reaction of financial statement users to published financial reports, including stock price reaction to the content of financial reports and the accuracy of analysts' forecasts of earnings and other performance measures. The quality of audited financial statements should be increased with the quality of the audit process applied.

I argue above that auditor expertise and access to superior technology and methodology potentially improve auditor decision making throughout each audit process. Evidence from the literature suggests that industry specialist audit firms possess expertise and superior technological and methodological resources, which they utilise in various stages of the audit processes. As such, it follows that the clients of these specialist firms may produce higher-quality financial statements. This section reviews the empirical literature examining the relation between the proposed indicators of the quality of audit inputs (large and industry specialist audit firms) and observed financial reporting outcomes (disclosure quality, accrual manipulations, market reactions to earnings surprises and analyst forecast accuracy).[15]

2.7.1 Disclosure Quality

Disclosure quality encompasses the quantity and decision usefulness of client disclosures additional to the data presented on the face of their financial statements. Such disclosures might include management forecasts of earnings or sales, information on shareholding breakdown, information on new products or business analysis. While I do not study disclosure quality directly, superior disclosure quality may affect the usefulness of audited financial statements for predicting future earnings. Existing research generates mixed evidence as to whether industry specialisation positively influences disclosure quality. Industry specialist audit firms

---

[15] Accounting conservatism is also considered to indicate financial reporting quality (Knechel et al. 2013). The traditional definition of accounting conservatism is a principle that anticipates no profit, but anticipates all losses (Bliss 1924, in Basu 1997). In the empirical literature, accounting conservatism is interpreted as reflecting 'the accountant's tendency to require a higher degree of verification to recognize good news as gains than to recognize bad news as losses' (Basu 1997, 7). Prior studies find that audit firm size and industry specialisation increase the levels of accounting conservatism (Basu et al. 2001; Krishnan 2005). Since accounting conservatism is not directly related to the 'valuation' perspective of financial reporting quality that I employ in my thesis, it is not explicitly discussed.

possess greater industry-specific expertise and can apply this to assist clients in developing and disseminating disclosures. Further, clients hiring these audit firms may demonstrate an intention to provide enhanced disclosures, as these firms are more likely to discover any deficiencies in client reporting (Dunn and Mayhew 2004). Therefore, it appears logical to expect that disclosure quality will be greater for clients of industry specialists. Consistent with this expectation, Dunn and Mayhew (2004) report a positive association between industry specialisation and client disclosure quality, as measured by analysts' rankings of company disclosures reported by the Association for Investment Management and Research, in unregulated industries. However, other research argues that industry specialist firms (as a proxy for high-quality audit firms) are more likely to be hired by clients who wish to disclose less and seek additional auditor credibility as a substitute (Peters et al. 2001). Under this argument, high-quality audit firms are regarded as a reputable intermediary that provides additional assurance to alleviate the effects of information asymmetry. Consistent with this contention, Peters et al. (2001) find that clients of industry specialist firms make less extensive commodity derivative disclosures than do clients of non-specialist audit firms.

Two studies have examined the relationship between audit firm size and management's forecast errors. Davidson and Neu (1993) predict and document a positive relation between management's forecast errors and audit firm size (as a proxy for high audit quality). Davidson and Neu (1993) assert that auditors are not responsible for clients' forecasts or forecast accuracy, and attribute their results to larger audit firms' influence on management's reporting discretion on earnings.[16] Conversely, Clarkson (2000) revisits Davidson and Neu's (1993) assumption and argues that high-quality auditors directly influence management forecasts. Clarkson contends that high-quality auditors pay careful attention to all aspects of a prospectus, including the forecasts, rather than narrowly concentrating on the financial statements. He predicts a negative relationship between management's forecast errors and audit firm size. Examining the one-year-ahead management earnings forecasts included in initial public offerings prospectuses, Clarkson (2000) finds that audit firm size (as a proxy for high-quality audit services) is associated with smaller management forecast errors after controlling for firm business risk and examining periods during which audit firms are responsible for audit-level assurance relative to review-level assurance.

2.7.2 Accrual Manipulations

A number of studies argue that high-quality auditors discipline managerial attempts to manipulate accruals and policy choices. Typically, these manipulations are measured by 'discretionary accruals', which represent the proportion of a firm's total accruals that cannot be explained by the firms' economic fundamentals. Measures of discretionary accruals are argued to capture management's discretionary judgment in financial reporting, which is thus commonly used to measure the degree of earnings management (Jones 1991; Healy and Wahlen 1999).[17] Several scholars argue that high-quality auditors constrain management attempts to manipulate accrual estimates

[16] There has been a plethora of evidence suggesting that high-quality audit service is a mechanism that constrains management's discretion regarding accounting estimates and policy choices, leading to decreased earnings management and increased earnings quality (Reichelt and Wang 2010; Francis and Yu 2009). This aspect of audit outcome is further elaborated in Section 2.7.2.

[17] Subramanyam (1996) proposes that discretionary accruals could be used as a signalling tool in cases where management smooths earnings. Such discretionary accruals may convey private information about a firm that may not usually be reflected under the historical cost accounting method.

and policy choices (Becker et al. 1998; Krishnan 2003; Payne 2008; Francis and Yu 2009), improving the reliability, persistence and predictability of accruals and published earnings (Sloan 1996; Xie 2001).

Industry specialists have greater expertise relevant to clients' operations and financial reporting concerns (Bedard and Chi 1993). They have also developed their reputation over time and have a greater incentive to protect their reputation to maintain clients and continue to earn fee premia (Craswell et al. 1995; Abbott et al. 2005). Therefore, specialist auditors are more likely to make better judgments throughout the audit processes (Taylor 2000; Low 2004). Specifically, while assessing clients' discretion on reported earnings and when disputing clients' accrual estimates and policy choices, specialist auditors are more likely to resist client pressure and negotiate with clients to reduce the incidence of manipulation (Kwon 1996; Gibbins et al. 2001, 2003). For these reasons, industry specialists are argued to reduce discretionary accruals, as a proxy for earnings management (Balsam et al. 2003; Krishnan 2003; Reichelt and Wang 2010).

Consistent with the reasoning above, there is empirical evidence that clients of audit industry experts have smaller absolute discretionary accruals, suggesting that industry specialists are better able to constrain earnings management, and thus improve earnings quality (Krishnan 2003; Balsam et al. 2003). Reichelt and Wang (2010) differentiate specialisation according to the level of geographic aggregation, defining national and city-specific industry specialists. [18] They find lower

[18] It has been argued that audit firms' industry expertise is largely city-specific because most auditors continue to service clients predominantly in one locale and derive deep industry expertise primarily out of offices in that locale (Ferguson et al. 2003; Francis et al. 2005b). More discussion on the city-specific measure of industry specialisation is provided in Chapter 4.

discretionary accruals among clients audited by city-specific industry specialists, and the lowest discretionary accruals for clients audited by *both* national and city-specific industry specialists. Adopting a less direct approach, Payne (2008) argues that industry specialist audit firms may constrain the earnings management behaviour associated with 'benchmark-beating' incentives,[19] which may in turn increase realised short-horizon forecast errors (i.e. realised errors in forecasts that are made immediately prior to the earnings announcement).[20] Using a sample restricted to Big N clients and an array of industry specialisation measures, Payne (2008) finds that industry specialisation increases short-horizon absolute forecast errors (i.e. decreases forecast accuracy) and reduces the likelihood of meeting or beating the forecast benchmark. However, using audit firm industry market share as a proxy for industry specialisation and a PSM approach, Minutti-Meza (2013) finds no evidence that audit firm specialisation is associated with a reduction in clients' propensity to meet or beat analysts' forecasts.

The related literature using audit firm size as a proxy for audit quality develops similar arguments and predictions regarding the relationship between audit firm size and clients' earnings management. Early research reports that clients of Big N auditors (as a proxy for large audit firms) have lower absolute discretionary accruals than do other firms (Becker et al. 1998; Francis et al. 1999). Bauwhede et al. (2000)

---

[19] Extensive prior research indicates that management has an incentive to manipulate earnings to 'meet or just beat' analysts' earnings forecasts to avoid negative shocks to companies' stock prices (Skinner and Sloan 2002; Dechow et al. 2003). For example, Bannister and Newman (1996) and Burgstahler and Eames (2006) show that management exercises discretion to alter accruals to 'meet or just beat' analysts' consensus forecasts.

[20] Forecast errors are the undeflated differences between actual earnings and forecast earnings. However, Payne's argument relies on the assumption that management attempts to 'meet' as opposed to 'beat' forecasts. If, for example, the pre-managed earnings forecast error is -\$0.1, reported earnings are manipulated upwards to \$0.15 (a case of 'just beating' analysts' expectations), with earnings forecast error increasing with earnings management behaviour. If an industry specialist constrains this behaviour, the forecast error should decrease.

provide further support for these findings in their study of a matched sample of public and private Belgian firms, which finds that the presence of Big N auditors suppressed income-decreasing earnings management. Francis and Yu (2009) argue and demonstrate further that Big N firms' office size is systematically associated with lower discretionary accruals.

Although academics frequently use discretionary accruals as an earnings quality proxy, the empirical meaningfulness of discretionary accrual measures is highly questionable (Ball 2013). The measurement of discretionary accruals is likely to be contaminated by non-discretionary accruals, and this contamination may induce a systematic bias in regression coefficients (Dechow et al. 1998; Ball 2013). Ball (2013) argues that stochastic shocks to business transactions are likely to be measured as 'discretionary accruals', and that controlling for these shocks in empirical models is problematic. Moreover, Francis (2011) claims that the extreme values of earnings in the statistical distribution are not necessarily an indicator of the misstatement of financial information; rather, they may reflect the cross-sectional variation in the statistical properties of earnings.

2.7.3 Market Responsive to Earnings Surprises

Stock market reaction to earnings announcements reflects whether financial information is useful for investors to make rational decisions regarding the value of a client. Stock prices are a function of the present value of expected future dividends, information regarding which is provided by accounting earnings (Ohlson 1990). Better quality earnings are argued to improve the reliability of information about a firm's future economic profits and dividend-paying ability (Kothari 2001; Teo and

Wong 1993). Therefore, stock prices or changes in stock prices may reflect the extent to which the market perceives the quality of earnings or a change in the quality of earnings. Clients' earnings response coefficient (ERC) and cumulative abnormal returns (CAR) are commonly used in the literature to assess the stock market responsive to earnings surprises. Since auditors provide assurance on the reported accounting numbers, the ERC or CAR measure can be used to assess investors' perceptions of auditors' reliability and financial information credibility and their subsequent decisions upon the perceptions. As large firms or industry specialist audit firms are generally more experienced and knowledgeable in the industries, they are more likely to influence clients to produce informative financial statements. As such, a positive link between clients of these audit firms and clients' ERC is expected (Teoh and Wong 1993; Balsam et al. 2003). Teoh and Wong (1993) report evidence that the ERC of companies audited by Big 8 (as a proxy for audit firm size) is greater than that of non-Big 8 clients. Balsam et al. (2003) find similar results by examining industry specialist firms. Similarly, if clients switch from a higher-quality audit firm to a lower-quality audit firm, a negative market reaction should be observed. Knechel et al. (2007) examine 159 firms that have switched between Big 4 auditors and find that clients experience the largest negative market reaction (as measured by the three-day CAR around the date of auditor switch) when clients switch from an industry specialist to a non-specialist audit firm. This negative association provides further evidence that audit quality increases the market's perception of earnings quality.

2.7.4 Analyst Forecast Accuracy

Analyst forecast accuracy is the absolute difference between analysts' forecasts of client earnings and the realisation of those earnings. Analyst forecast accuracy is perhaps the most direct way to assess whether the primary objective of financial reporting quality (and thus audit quality), which emphasises the importance of earnings for users' prediction of firms' future performance, has been achieved. Recall that large or industry specialist audit firms may constrain management attempts to manipulate accrual estimates and policy choices, improving the reliability of accruals and thus earnings (Krishnan 2003; Balsam et al. 2003; Reichelt and Wang 2010; Francis et al. 1999). There is substantial empirical evidence showing a positive relationship between accrual quality and analyst forecast accuracy (Abarbanell and Lehavy 2003; Ahmed et al. 2005). Therefore, if superior audit quality improves the quality of clients' financial reports, this should increase the accuracy of analysts' earnings forecasts. Below I describe the empirical evidence regarding the association between accrual quality and analyst forecast accuracy, followed by a discussion of the literature linking audit firm industry specialisation to forecast accuracy.

Several papers examine the relationship between accrual quality and analyst forecast accuracy, suggesting that discretionary accruals in general mislead analysts unless there is an event signalling the presence of earnings management. Abarbanell and Lehavy (2003) find that extreme forecast errors are positively correlated with the presence of large discretionary accruals in the realised earnings figures, suggesting that analysts are not able to anticipate perfectly firms' earnings management behaviour. Ahmed et al. (2005) find that analyst forecast accuracy is negatively associated with firms' prior period discretionary accruals, suggesting that analysts either do not recognise the lower persistence of discretionary accrual component of earnings or are not motivated to do so. However, Wilson and Wu (2011) show that the sensitivity of analysts' forecast errors to current-period discretionary accruals is lower in cases where a clear public signal of earnings management incentives exists, consistent with analysts' recognition of client attempts to bias earnings.

Recent empirical research theoretically connects high-quality audit services with greater predictability of earnings and tests this relationship using analysts' forecast errors.[21] BCK (2008) argue that higher-quality auditors subtend higher earnings quality and thus decrease forecasting task complexity. Analysts are thus more likely to evaluate the implications of current earnings information correctly, and in turn make earnings forecasts that are more accurate. Therefore, they predict a negative (positive) relation between absolute forecast errors (accuracy) and audit quality. BCK (2008) find that analysts' forecast errors are lower (i.e. forecasts are more accurate) for clients of Big N audit firms, and that audit firm industry specialisation reduces forecast errors, but only for clients of non-Big N auditors. However, Lawrence et al. (2011) show that there is no relationship between analysts' forecast errors and audit firm size after matching client characteristics for samples of fundamentally similar client firms.

The theory underpinning BCK's findings that analyst forecast accuracy positively correlates with high-quality audit services has been examined deeply in two recent

[21] As noted in Section 2.7.2, Payne (2008) reports that auditor industry specialisation is negatively related to forecast accuracy. However, Payne's study does not explicitly examine the importance of audit quality on users' prediction of firms' future earnings. Rather, it focuses on its impact on managerial benchmark-beating behaviour, and thus earning quality. Short-horizon forecast accuracy is used to assess this earnings quality.

papers. Choi and Kwon (2008) argue analytically and use a theoretical prediction model to show that accounting errors decrease with high-quality audit, which then enhances analysts' ability to predict firms' future earnings. These arguments are consistent with BCK's results. He et al. (2011) investigate the influence of audit quality on analysts' earnings forecasts by focusing on analysts' information environment. Using Barron et al.'s (1998) model, He et al. (2011) find that high-quality audits (proxied by either audit firm size or industry specialisation) are associated with a greater use of common information by analysts when making forecasts, and with greater precision of analysts' common *and* private information. While these studies attempt to explain the rationale behind BCK's findings, they are silent about Payne's predictions and findings. [22] In Chapter 5, I explore the differences in research design across the BCK and Payne studies and show that their contradictory findings derive from model specification and the choice of deflators for key variables.

2.8 Chapter Summary

In this chapter, I reviewed the literature relevant to the connection between audit firm industry specialisation and measures of audit or financial reporting quality. I maintained that audit quality reflects the auditor's impact on overall financial reporting quality, and that audit quality can be inferred by observing financial reporting outcomes (Section 2.2). A theoretical framework adapted from that of KKPSV, which I use to describe the means by which the supply of higher-quality audit services may affect financial reporting quality, was described in Section 2.3. I

[22] In the footnote, He et al. (2011, 3) suggest that Payne 'ignores the direct role that financial reporting plays in developing earnings forecasts (as distinct from measuring the error)'. No further discussion is provided with respect to the inconsistent findings generated by Payne (2008) and BCK (2008).

next defined and described key audit inputs and examined their importance in Section 2.4, with an emphasis on industry specialist auditors' superior quality of audit inputs. In Section 2.5, I analysed the empirical evidence concerning the relationship between various indicators of audit input quality and the quality of audit processes, again emphasising evidence of the role of audit firm industry specialisation. The impact of auditors' private incentives on the extent to which the quality of audit inputs flow through to superior process performance was examined in Section 2.6. Finally, Section 2.7 examined the interrelationship between audit inputs, audit processes and audit outcomes. Audit outcomes are used to assess financial reporting quality, and thus audit quality. Of the various audit outcomes identified, I argue that analyst forecast accuracy is the most direct measure of whether the primary objective of financial reporting quality (thus audit quality) has been achieved. Flowing from this logic, I develop hypotheses regarding the generic and cross-sectional relationships between audit firm industry specialisation and analyst forecast accuracy in Chapter 3.

CHAPTER 3: HYPOTHESIS DEVELOPMENT

3.1 Introduction

In Chapter 2, I reviewed the extant research examining the means by which the supply of higher-quality audit services may result in superior financial reporting quality, with an emphasis on the role of audit firm industry specialisation in that context. I argued that audit quality could be inferred from the properties of the financial reports published, and from the impact of the financial reports upon markets and market participants. In this chapter, I develop hypotheses concerning the relationship between audit firm industry specialisation and the performance of a key group of market participants: securities analysts. First, I analyse the predictions and findings of prior studies that have examined the impact of audit firm industry specialisation on the accuracy of analysts' forecasts, and develop predictions regarding the generic relationship between these variables in Section 3.1 (Hypotheses 1a and 1b). Although tests of the overall relationship between audit firm industry specialisation and forecast accuracy may produce evidence consistent with a causal relationship, I develop additional hypotheses, tests of which may present more convincing evidence regarding the existence of a causal effect of audit quality on analyst forecast accuracy. These additional hypotheses predict that the relationship between audit quality and analyst forecast accuracy varies cross-sectionally with the inherent difficulty of the forecasting task. Hypothesis 2 examines the association between audit quality and analyst forecast accuracy when the client firm's operating risk varies (Section 3.2), while Hypotheses 3a and 3b explore this association conditional on the estimated quality (expertise) of the analysts following a client firm (Section 3.3). Section 3.4 concludes the chapter.

3.2 Hypotheses 1a and 1b—Audit Firm Industry Specialisation, Forecast Accuracy and Forecast Horizon

The objective of financial reporting is to provide information that is useful to financially literate investors, lenders and creditors when assessing the prospects for future net cash flow (SFAC No. 8, FASB 2010 OB2 and OB3). Measuring and reporting earnings and its components is the primary focus of financial reporting because this information is central to the prediction of earnings and future cash flows (SFAC No. 8, FASB 2010 BC1.31; OB 17; OB18). Therefore, the usefulness of financial reporting is a function of earnings quality, which in turn is heavily influenced by the quality of accrual estimates and supporting disclosures. Accrual quality reflects the extent to which accruals shift or adjust the realisation of cash flows, such that earnings can better capture the firm's underlying performance (Dechow and Dichev 2002) and the quality of accruals decreases in the magnitude of estimation errors (whether intentional or not) (Allen et al. 2013; Dechow and Dichev 2002). Disclosure quality embraces the quantity and decision usefulness of client disclosures (e.g. management forecast, and supplemental information that enhances the interpretation of the performance figures). Thus, when clients' accrual and disclosure quality are high, published financial reports should be more useful for predicting future earnings by financially literate users such as securities analysts.

Prior studies argue and show that clients have better financial reporting outcomes as reflected in accrual and disclosure quality when the financial reports are audited by a high-quality audit provider, typically proxied by audit firm size and industry specialisation (Becker et al. 1998; Krishnan 2003; Dunn and Mayhew 2004; Clarkson 2000). These superior reporting outcomes are argued to derive from auditors' domain-specific expertise relevant to clients' operations and their strong incentive to protect their reputation capital (Bedard and Chi 1993; Solomon et al. 1999; DeAngelo 1981). Reputation concerns motivate auditors to improve their judgment quality and negotiate with clients over the application of accounting principles (DeAngelo 1981; Gibbins et al. 2001, 2003), while auditors' greater expertise provides them with greater confidence in their propositions and improves their judgment and decision making quality throughout the auditing processes (Gibbins et al. 2001, 2003; Kwon 1996; Low 2004). For instance, high-quality auditors may be more likely to resist management's pressure over accrual estimates and policy choices and negotiate with management to reduce the incidence of manipulations (Kwon 1996). Their clients have also been found to have lower discretionary accruals (Becker et al. 1998; Krishnan 2003; Balsam et al. 2003). Similarly, high-quality auditors may reduce unintentional estimation errors in clients' reported earnings, reducing the difference between the reported earnings and the 'true' economic earnings (Watkins et al. 2004; Choi and Kwon 2008). If superior audit quality is associated with lower discretionary accruals and unintentional estimation errors, it should increase the quality of accruals, which in turn improves the reliability of earnings for users' prediction of future performance (BCK 2008). Further, high-quality auditors may improve the quality and breadth of supporting disclosures (Dunn and Mayhew 2004; Clarkson 2000), which in turn may be of use in forecasting clients' future earnings. On the basis of the argument and empirical findings discussed above, I expect that the financial reports of clients audited by a high-quality audit provider should be more useful for predicting future earnings, and that this greater usefulness will be reflected in the accuracy of analysts' earnings forecasts.

Notwithstanding the above argument, there are reasons that the empirically observed relationship between audit quality and forecast accuracy may be of opposite direction to that suggested above. Payne (2008) argues that managers perceive a future private benefit in 'meeting or just beating' the level of earnings implied by end-of-year consensus forecasts, [23] and may attempt to manipulate accruals to achieve these benchmarks. High-quality audit providers may be more effective in constraining clients' attempts to manage current earnings towards the zero forecast error level. As such, high audit quality may be associated with larger forecast errors (i.e. lower forecast accuracy). Thus, the empirically observed relationship between audit quality and forecast accuracy is likely to depend on the intra-year timing of the forecasts studied.

Prior studies examining the relationship between audit quality and analysts' forecast errors test their propositions by focusing on the accuracy of analysts' short-horizon forecasts (Payne 2008; BCK 2008). Short-horizon forecasts (also known as 'end-of-year forecasts') are the forecasts outstanding at the client firms' reporting date, and are typically issued or revised in the weeks immediately prior to reporting. While BCK (2008) propose that high-quality audit providers improve analyst forecast accuracy, they find a positive association between forecast accuracy and the presence

[23] Burgstahler and Eames (2006) provide visual and statistical evidence that management's avoidance of negative earnings surprises becomes progressively weaker when forecasts are issued from 271 to 360 days before the current year earnings reporting date (long-horizon forecasts in my study) relative to when forecasts are issued close to the earnings release date (short-horizon forecasts).

of Big N audit firms, but no significant association between Big N audit firms' portfolio-share industry specialisation and analyst forecast accuracy. Payne (2008) argues and shows that industry specialist auditors reduce the effectiveness of clients' attempts to manage current earnings towards consensus forecasts, which in turn increase analysts' absolute forecast errors (lower forecast accuracy).[24]

I argue that the relationship between short-horizon forecast errors and the provision of high-quality audit services is potentially confounded by the competing effects described above. On the one hand, industry specialist auditors may improve the usefulness of prior period or interim financial reports for predicting earnings, and to the extent that the current auditor was responsible for those prior period reports, a negative association between absolute forecast errors and audit quality may be expected. However, short-horizon forecasts are likely to be the focus of benchmark-beating incentives, and thus a superior auditor may be more likely to constrain managerial attempts to bias current-period earnings towards market expectations. If both of these effects occur simultaneously, the directional impact (if any) of audit firm industry specialisation on analysts' absolute forecast errors is unclear. Thus, I state the following non-directional maintained hypothesis regarding the association between audit firm industry specialisation and analysts' short-horizon forecast errors:

H1a: *Analysts' short-horizon absolute forecast errors are associated with audit firm industry specialisation.*

---

[24] However, a recent study finds no evidence that audit firm specialisation, measured using market-share approach, is associated with a reduction in clients' propensity to 'meet or just beat' analysts' forecasts once client characteristics are appropriately controlled (Minutti-Meza 2013).

As the competing theoretical impacts of audit firm industry specialisation on short-horizon forecast errors impair the interpretation of any tests of that relationship, I propose a more direct test focusing on the impact of industry specialisation on analyst long-horizon (beginning-of-year) absolute forecast errors. Long-horizon forecasts are those issued immediately after the release of prior period earnings. I argue that their accuracy is a more powerful meter of audit quality for two reasons. First, over this forecast horizon, audited financial reports represent a relatively large proportion of the information available for predicting future earnings; thus, the accuracy of these forecasts is logically more sensitive to variations in the quality of the audited reports.[25] Second, with long-horizon forecasts, there is a lesser likelihood of earnings management behaviour aimed at 'meeting or just beating' these forecasts than is the case with short-horizon forecasts. Put simply, while there are well-known incentives for firms to attempt to 'meet or just beat' earnings forecasts current at the time earnings are released (Burgstahler and Eames 2006), there is no obvious incentive for a firm to manipulate earnings to achieve a redundant target. Therefore, I argue that long-horizon forecast accuracy is a more direct and less noisy measure of the extent to which the objectives of financial reporting are being satisfied.

To illustrate the lesser likelihood of 'benchmark-beating' (or expectation management) behaviour with respect to long-horizon forecasts, I examined the proportions of forecast errors for which actual earnings meet or beat consensus by 1 cent or less. In the case of short-horizon forecast errors, approximately 20 per cent of all forecast errors fall in this range, whereas less than 3 per cent of long-horizon

[25] He et al. (2011) study forecasts issued immediately after the release of prior year earnings and argue that these forecasts are most directly affected by the quality of published accounting information.

forecast errors are between 0 and 1 cent.[26] This suggests that earnings management to avoid negative earnings surprise are substantially weaker for long-horizon forecasts (which are issued an average of 304 days ahead of the release of actual earnings). Thus, the impact of higher audit quality on the reliability of prior year earnings is likely to dominate any association between audit quality and 'benchmark-beating' behaviour. Thus, I predict that clients of industry specialist auditors have lower long-horizon absolute forecast errors:

***H1b:** Analysts' long-horizon absolute forecast errors are negatively associated with audit firm industry specialisation.*

3.3 Hypothesis 2—Audit Firm Industry Specialisation, Client Firm Operating Risk and Forecast Accuracy

Hypothesis 1b predicts that high-quality audit services increase the usefulness of published financial reports for forecasting future earnings and thus improve the accuracy of analysts' long-horizon forecasts. To further examine whether any observed empirical relationship between audit quality and forecast accuracy may be causal, I develop additional hypotheses that predict cross-sectional variation in the association between auditor industry specialisation and forecast accuracy. One source of predicted variation in this relationship is client firm operating risk. In this section, I develop a hypothesis in which I argue that where clients' operations are inherently very stable, and accrual estimates are relatively easy to verify, audit

[26] To further examine the association between audit firm industry specialisation and benchmark beating, I estimate probit regressions similar in form to Payne (2008), to model the probability of a client firm achieving a 0 or 1 cent forecast error (see Appendix B). I show that, for short-horizon forecasts, audit firm industry specialisation is significant negatively associated with the likelihood of 'meeting or just beating' consensus. However, when applied to long-horizon forecasts, there is no relation between such benchmark-beating behaviour and audit quality.

quality has a reduced influence on the accuracy of earnings forecasts, and therefore indicators of audit quality, such as auditor industry specialisation, should be more strongly associated with forecast accuracy. Testing this hypothesis has the potential to bring stronger evidence to bear regarding whether any findings from tests of H1b are consistent with the existence of a causal relationship.

My thesis defines client firm operating risk as subsuming the uncertainty arising from business and financing activities. While this risk is not directly caused by accounting decisions and estimates, its presence may decrease the precision of information available for the prediction of firms' future cash flows and earnings (Minton et al. 2002).[27] The information critical to the prediction of firms' future performance is the accrual component of earnings (Barth et al. 2001; Minton et al. 2002). Estimation errors in accruals (intentional or otherwise) may affect the extent to which accruals map to future cash flows, thus decreasing the extent to which the reported earnings are useful for predicting future earnings (Allen et al. 2013).[28] Empirical evidence shows that analysts' forecasts do not fully incorporate the lower persistence of accrual information (Ahmed et al. 2005; Abarbanell and Lehavy 2003). Therefore, client firm operating risk is likely to be associated with greater accrual estimation errors, and thus greater absolute earnings forecast errors.

---

[27] This definition of operating risk is closely related to auditors' concept of inherent risk, but differs in the fact that it refers only to externally observable measures of variability in firms' performance. The traditional audit concept of inherent risk includes these factors, but may also be affected by other factors such as prior audit results, the auditor–client-specific experience, the existence of related parties and non-routine transactions (Arens et al. 2013, 239).

[28] Allen et al. (2013) decompose accruals into 'good accruals' and 'accrual estimation errors'. 'Good' accruals reverse when the realisation of cash flow is anticipated, so that actual realisation of cash flow does not affect earnings. 'Accrual estimation errors' do not anticipate future benefits and their reversal are not offset by the anticipated cash flow, resulting in an impact on earnings. Allen et al. (2013) argue and find that the low persistence of earnings is caused by the 'accrual estimation error', not by 'good' accruals.

I argue that audit quality should more greatly affect analysts' absolute forecast errors when client operating risk is high. This prediction reflects assumed cross-sectional differences in both the difficulty of the audit task and the level of audit effort supplied, which I now describe. Recall that high-quality audit providers have been found to reduce clients' accrual estimation errors to a greater extent than do other audit providers (Balsam et al. 2003; Krishnan 2003). This, in turn, improves the reliability of clients' published financial statements as a basis for predicting their future earnings. The impact of audit quality on accrual estimation errors, and thus earnings predictability, should be greater when the scope for such errors is larger, and the scope for accrual estimation errors is greater when the client operating risk is higher, as discussed above. It logically follows that if audit firm industry specialisation is associated with superior audit quality, the extent of specialisation should have a greater effect in improving earnings predictability (decreasing absolute earnings forecast errors) when client firm operating risk is higher. For example, if a client has zero operating risk, accrual estimates should be perfect and earnings can be predicted without error, implying that industry specialist auditors should have no role in constraining accrual estimation errors and improving earnings forecast accuracy. However, when a client has a high level of operating risk, accrual estimation errors are higher and earnings forecast for this client will be less accurate. For these riskier clients, industry specialist auditors are likely to reduce clients' accrual estimation errors to a greater extent than would other audit providers, improving the usefulness of financial information for analysts' predictions of earnings. Thus, the impact of audit industry specialisation in improving earnings forecast accuracy is greater when the client operating risk is higher.

Second, studies of the closely related concept of inherent risk find that this measure of risk is positively associated with audit effort, and that returns to the additional effort are greater for high-quality providers than for other auditors (O'Keefe et al. 1994; Schelleman and Knechel 2010; Caramanis and Lennox 2008). Auditors increase their effort with assessed inherent risk (O'Keefe et al. 1994), and this greater audit effort has been shown to constrain income-increasing earnings management (i.e. decrease accrual estimation errors) (Caramanis and Lennox 2008). Prior studies also show that high-quality audit providers exert greater effort in response to signals of higher risk in clients' operations (e.g. higher levels of short-term accruals) than other auditors, and that the returns to each additional unit of effort are greater for these high-quality auditors (Schelleman and Knechel 2010; Caramanis and Lennox 2008). Therefore, when risk is higher, industry specialist auditors are predicted to supply greater effort than non-specialists, and the additional audit effort supplied is likely to have a greater impact in reducing the errors in accruals (and thus earnings), which in turn increases the predictability of earnings. Based on the above rationale, I argue that if any observed relationship between auditor industry specialisation and long-horizon absolute forecast errors is causal, audit firm industry specialisation should decrease absolute forecast errors to a greater extent when the riskiness of clients' operations is higher. Hypothesis 2 reflects this prediction:

H2: *The negative association between analysts' long-horizon absolute forecast errors and audit firm industry specialisation increases with the level of the client firm's operating risk.*

3.4 Hypotheses 3a and 3b—Audit Firm Industry Specialisation, Analyst Quality and Forecast Accuracy

In this section, I develop additional hypotheses relating to the predicted cross-sectional variation in the impact of audit quality on analyst forecast accuracy. Hypotheses 3a and 3b focus on differences in the quality (expertise) of the analysts issuing the forecasts of a client firm's earnings, and the extent to which the accuracy of those forecasts is affected by industry specialist auditors. These hypotheses predict that audit firm industry specialisation has a greater impact on the accuracy of lower-quality analysts, and they imply tests that may provide further evidence supporting (or contradicting) the existence of a causal relationship between audit quality and analyst forecast accuracy. Tests of H3a and H3b also directly examine whether audit quality contributes to the extent to which the objectives of financial reporting are met. The Conceptual Framework states that an objective of general-purpose financial reporting is to provide financial information relevant to decision making that 'will meet the needs of the maximum number of primary users' (SFAC No. 8, FASB 2010 OB8), where primary users refer to existing and potential investors (including analysts), lenders and other creditors (SFAC No. 8, FASB 2010 OB2 and BC1.9). Although the Conceptual Framework assumes that primary users (including analysts) possess a minimum level of financial competence, it logically follows that financial reports are of superior quality when they are useful to a greater number of financially competent users.

Consistent with much of the analyst literature, my thesis assumes that analyst quality derives from individual analysts' access to information outside that contained in the financial statements,[29] and from the accretion of superior ability to identify and process relevant complex information.[30] Since analyst quality cannot be observed directly, prior studies employ numerous proxies for analyst quality, including analysts' general and firm-specific forecasting experience, the size of the analysts' employer, analysts' prior forecast accuracy, the simplicity of the analysts' portfolios of covered firms and 'All-Star' analyst status. There is empirical evidence of a positive relationship between these proxies for analyst quality and analyst forecast accuracy (Clement 1999; Clement and Tse 2005; Clement 2007; Kim et al. 2011; Drake and Myers 2011).[31] I discuss the literature relevant to the selection of analyst quality proxies in Chapter 4.

To the extent that the superiority of high-quality analysts reflects their superior access to private information and ability to identify and analyse both public and private information, higher-quality analysts will be less reliant on the quality of published financial reports when making predictions of future earnings. For example, high-quality analysts have been shown to adjust forecasts according to the level of total accruals present (Drake and Myers 2011). Similarly, high-quality analysts may

---

[29] I contend that this information is not restricted to information obtained legally or otherwise from corporate insiders and may extend to other relevant non-earnings information signals identified by analysts.

[30] The information effect reflected in analyst quality measures (i.e. access to more private information) may have become weaker after the introduction of Regulation Fair Disclosure (Reg FD) on 23 October 2000. Reg FD prohibits analysts' access to private information from managers; thus, one of the achievements of Reg FD has been a reduction in the information asymmetry among all analysts (Eleswarapu et al. 2004). Previous evidence shows that analyst forecast accuracy is significantly different when comparing the pre-Reg FD and post-Reg FD periods (Heflin et al. 2003; Findlay and Mathew 2006). This information argument about analyst quality may be relevant to my study, as my sample (which is described in Chapter 6) starts from year 1989. Therefore, to year 2000, analysts' private information may have influenced their forecast accuracy. In view of this, I investigate the impact of Reg FD on my hypothesised relationships in my sensitivity tests.

[31] Unlike most prior analyst quality studies, which examine short-horizon forecast accuracy, my thesis focuses on long-horizon forecasts, which are made immediately after the release of previous years' financial reports. The only extant study in this area is Drake and Myers's (2011) investigation of whether analyst quality is related to analysts' long-horizon forecast accrual-related over-optimism. My thesis explicitly states that analyst quality increases forecast accuracy, and I examine whether analyst quality moderates the relation between audit quality and forecast accuracy.

use information obtained from other sources to predict earnings and improve forecast accuracy when the precision of publicly available information (i.e. financial reports) is low (Keskek et al. 2013). Conversely, lower-quality analysts may benefit relatively greatly from improvements in financial reporting quality (and thus audit quality) because these analysts are relatively dependent on primary information sources such as general-purpose financial reports, and may be less able to undo distortions associated with abnormal accruals. It thus follows that if higher-quality audit providers, such as industry specialist auditors, improve the usefulness of published financial reports for predicting future earnings, the resulting impact on forecast accuracy should be greater for lower-quality analysts.

From this general contention, I develop two specific hypotheses pertaining to the association between auditor industry specialisation, analyst quality and forecast accuracy, which differ according to the unit of analysis (i.e. firm-years or analyst-firm-years). First, I argue that the *average quality of the analysts covering a firm* in a given year moderates the association between audit firm industry specialisation and analysts' consensus long-horizon forecast errors. Where average analyst quality is high, their aggregate information set is also relatively high, and accordingly the importance of basic signals such as those contained in the published financial reports is relatively low. Conversely, where average analyst quality is low, analysts may rely more heavily on easily accessible public information, such as that contained in the published financial statements, when making earnings forecasts. In these cases, I argue that the quality of those financial statements (and thus audit firm industry specialisation) should have a greater influence on analysts' absolute forecast errors. Thus:

H3a: *The negative association between analysts' long-horizon absolute forecast errors and audit firm industry specialisation decreases with the average quality of analysts following a firm.*

I apply similar logic in developing a hypothesis concerning the relative forecast accuracy of *individual* analysts following a given client firm. To the extent that greater audit firm industry specialisation leads to higher-quality financial reports and has a greater impact on the forecast accuracy of lower-quality analysts, the difference in forecast accuracy between the 'worst' and 'best' quality analysts following a firm should be reduced. Hypothesis 3b is therefore:

H3b: *Greater audit firm industry specialisation reduces the difference in analysts' absolute forecast errors between the 'worst' and 'best' quality analysts following a firm.*

3.5 Chapter Summary

This chapter developed three hypotheses in accordance with the research questions identified in Chapter 1. Hypotheses 1a and 1b concern the generic relationship between audit firm industry specialisation and analyst forecast accuracy. I then developed two hypotheses that I argue bring stronger evidence to bear concerning the possible existence of a causal relation between auditor industry specialisation and forecast accuracy. Hypothesis 2 predicts that audit firm industry specialisation is of greater importance on forecast accuracy when the firm's operating risk is higher. Hypotheses 3a and 3b propose that audit firm industry specialisation has a greater impact on the forecast accuracy of lower-quality analysts, improving the extent to which the financial reports are useful for a broader range of users rather than experts. In the next chapter, I discuss in detail the empirical proxies for my key measures: audit firm industry specialisation and analyst quality.

CHAPTER 4: EMPIRICAL PROXIES FOR AUDITOR AND ANALYST QUALITY

4.1 Introduction

I argued in the previous chapters that the empirical impact of the quality of audit services provided to a client could be inferred by observing audit outcomes, including the usefulness of the resulting financial reports for predicting future earnings. I also developed five hypotheses relating to the impact of audit firm industry specialisation on the predictability of future earnings, each of which implies a dependent variable that is a function of earnings predictability. To test each of my hypotheses, a measure of the extent to which the quality of audit services is expected to vary (audit firm industry specialisation) is required, and the tests of Hypotheses 3a and 3b also require metrics for 'analyst quality'. This chapter introduces and explains my choice of empirical proxies for each of these key constructs, which represent the test variables in my thesis. I introduce these key measures here, as the nature of the chosen measures affects subsequent modelling choices. The balance of this chapter is organised as follows. Section 4.1 discusses in detail the use of audit firm industry specialisation as a proxy for superior audit quality and examines the various extant measures of audit firm industry specialisation. In Section 4.2, I identify and explain the empirical proxies for analyst quality used in my study. Section 4.3 concludes the chapter.

4.2 Audit Firm Industry Specialisation and Audit Quality

My thesis focuses on whether audit firm industry specialisation improves the quality of the audit services supplied, which would in turn enhance the predictability of future earnings. As early as the late eighteenth century, Smith (1776) proposed that specialisation increases economic efficiency. Smith (1776) argues that the division of labour is central to economic efficiency, and he proposes a number of reasons for this effect, two of which are particularly relevant to the audit market. First, the division of labour reduces the scope of the work tasks required of particular labour suppliers, encouraging the accumulation of skills through repetition. Second, the division of labour encourages invention and the use of machinery dedicated to particular tasks, the application of which facilitates and abridges labour effort and time.[32] It is thus arguable that greater specialisation (division of labour) within the production of audit services may improve efficiency, for reasons elaborated upon below.

Large audit firms recognise the importance of industry specialisation, and since the late twentieth century have shown an increasing trend to concentrate their engagements in particular industries (Gramling and Store 2001; Hogan and Jeter 1999). For example, KPMG emphasises that 'we structure ourselves by industry sector as well as by our three core services of Audit, Tax and Advisory' (KPMG 2014), while PricewaterhouseCoopers states that they focus on audit and assurance, tax and consulting services and concentrate services in 16 key industries (PWC 2014). In addition, Ernst and Young use an industry-focused approach to provide audit and advisory services (EY 2014). Audit firms for which this concentration of

[32] Another benefit of specialisation proposed by Smith (1776) is that the division of labour saves time lost in moving from one task to another.

clients is particularly great are referred to as 'industry specialists'. Recall that industry specialists are considered to develop general and domain-specific knowledge and problem-solving skills (and thus expertise) relevant to their clients' industries (Bedard and Chi 1993). In addition, industry specialists invest disproportionately in technologies, physical facilities, personnel and organisation control systems in industries in which they choose to concentrate (Simunic and Stein 1987; Gramling et al. 2001). The accumulation of greater expertise and investment in particular industries enables industry specialist auditors to apply superior judgment and decision making throughout the audit processes, which may lead to better audit outcomes, such as the production of financial reports that are more useful for predicting future earnings.

The extant literature employs several empirical proxies for the existence and intensity of audit firm industry specialisation. In experimental and survey-based research, individual auditors' level of industry-specific audit experience is frequently used to proxy industry specialisation. Archival research typically measures audit firm industry specialisation by either the extent to which an audit firm's total audit fee revenue is earned within a particular industry (the 'portfolio-share' measure) or by the audit firm's share of aggregate audit fees paid by clients in a particular industry (the 'market-share' measure). The conceptual basis of these and other measures of industry specialisation and the extent to which they are likely to capture underlying audit quality are discussed in the following sub-sections.

4.2.1 Audit Firm Industry Specialisation—Market-share Measure

A number of studies use a market-share-type metric to capture the effect of industry specialisation on auditor expertise (Balsam et al. 1993; Craswell et al. 1995; Godfrey and Hamilton 2005). Market-share measures of industry specialisation reflect an audit firm's share of the total audit fee revenue generated in an industry. Due to data limitations, many academics use client total assets as a proxy for client fees (Payne 2008; BCK 2008; Gul et al. 2009).[33] The generic form of the continuous market-share measure of industry specialisation is shown below.[34]

Market share	=	the sum of the square root of the total assets of the clients that an audit firm services in a particular industry divided by the sum of the square root of the total assets of all clients in that industry

Market-share measures of industry specialisation are argued to capture expertise arising from two sources: knowledge spillovers and economies of scale. Yardley et al. (1992, 151) argue that industry market share is related to expertise because, *ceteris paribus*, greater market share increases the value of knowledge transfers between services provided to multiple clients. Audit firms with high industry market share may experience greater knowledge spillover benefits than other audit firms, because they supply a greater amount of audit services to fundamentally similar clients. Krishnan (2001) argues that industry expertise is associated with production efficiencies through economies of scale, which result in lower-cost audits.[35] An audit firm will concentrate operations in an industry until the production efficiencies through economies of scale are fully absorbed. Thus, the existence of economies of

[33] Audit fee data are only available from Audit Analytics for years after 1999. Many other studies use clients' total revenue (Krishnan 2003; Godfrey and Hamilton 2005) or the number of clients of an audit firm (Chin and Chi 2009; Balsam et al. 2003) to proxy audit fees and generate similar results as to the studies using clients' total assets.

[34] I describe the dichotomous measure in detail in Chapter 5.

[35] The underlying assumption to the argument is that homogeneous products, price competition and price-inelastic aggregate demand exist in the audit market (Krishnan 2001).

scale is argued to imply that industry expertise is reflected in audit firm industry market share (Krishnan 2001).

While the market-share measure of industry specialisation may have desirable attributes such as the ability to capture the value of knowledge spillovers and economies of scale, this approach has potential significant limitations. The first limitation stems from product differentiation theory, which assumes that audits are not homogeneous, and that differences in (perceived) audit provider quality affect the demand curve facing each auditor (Krishnan 2001). If industry expertise is a component of auditors' product differentiation strategy, greater expertise will increase the slope of the demand curve facing the audit firm. If cost functions are similar across expert and non-expert firms, Krishnan (2001, 131–132) shows that the expert auditor will have a *smaller* market share than the non-expert in equilibrium. Consequently, Krishnan (2001) concludes that the asserted positive association between industry expertise and market share cannot hold if expertise is a component of auditors' product differentiation strategy.

Second, the market-share measure of industry specialisation is increasing in market dominance, which in turn may imply well-known dysfunctional consequences (Yardley et al. 1992). In addition to the market power held by a single audit firm, the likelihood of a small number of (potentially colluding) firms dominating a market increases with firms' market share. As market power and/or market concentration reduce competition and may induce collusion, the leading ('industry specialist'') firms may experience lower returns to the provision of quality (Yardley et al. 1992). While regulators have expressed concerns regarding the negative impact of monopoly power and market concentration on the quality of audit outcomes (Great Britain: House of Lords 2010a, 2010b, 2010c; European Union: European Commission 2010; United States: General Accounting Office 2003a, 2008), academics have examined their consequences empirically (Francis et al. 2012; Boone et al. 2012). Francis et al. (2012) find that earnings quality is lower in countries with greater within-Big 4 market concentration, and attribute this to weaker competition among those audit firms. Boone et al. (2012) argue that audit firms with leading market power are likely to collude and devote less audit effort; they show that auditors' tolerance of earnings management (indication of lower audit quality) increases with audit market concentration.[36]

In summary, while greater market share may increase knowledge spillover benefits and economies of scale, the potentially confounding effect of market dominance and product differentiation effects suggest that the market-share measure of industry specialisation is at best a very noisy measure of the quality of audit services provided; it may even be associated with reduced quality.[37] Therefore, I limit the use of this proxy to my sensitivity analyses. An alternative measure of industry expertise, argued to be superior to the market-share approach, is examined in the next section.

---

[36] Boone et al. (2012) propose an alternative view that states that market dominance may reduce auditors' need to curry favour with clients through means of sacrificing some degree of scepticism because of fear of being replaced. Dominant audit firms are thus more likely to provide better audit services. However, this view is not supported by their results.

[37] The third limitation concerns the noisiness of the empirical measures of market share. Typically, the market-share approach only identifies large audit firms (Big N) as industry specialists because these firms tend to have large clients in a particular industry. However, knowledge spillovers may also occur in relatively small audit firms that concentrate activities on a few small clients in an industry. These firms may also develop expertise from providing audit services to a group of similar small firms, or servicing a particular client firm for many years in that industry. However, they are less likely to be recognised as specialists due to the relatively small size of their clients (Yardley et al. 1992; Minutti-Meza 2013). This limitation of the market-share measure has little impact on my study because my sample is restricted to clients of Big N firms.

4.2.2 Industry Specialisation—Portfolio-Share Measure

The audit firm portfolio-share measure of industry specialisation is a frequently employed alternative (or complement) to the market-share measure described above (e.g. Krishnan 2001, 2003; BCK 2008; Payne 2008). The portfolio-share industry specialisation measure attempts to capture the extent to which audit firms concentrate their productive activities (and thus fee base) within particular industries. This measure implicitly assumes that audits and audit quality are heterogeneous, and that industry specialisation may be one means through which auditors differentiate their product (Krishnan 2001). The generic estimation of an audit firms' portfolio-share measure is described below:[38]

Portfolio share	=	the sum of the square root of the total assets of the clients that an audit firm services in a particular industry divided by the sum of the square root of the total assets of all clients of that audit firm

The portfolio-share measure emphasises industry share within an audit firm's portfolio, rather than the audit firm's share within an industry. In this way, the measure focuses on auditor industry expertise resulting from the audit firm's strategic business decisions. Simunic and Stein (1987) and Neal and Riley (2004) argue that audit firms invest heavily in industry-specific technologies, physical facilities, personnel and organisation control systems in industries in which they are more likely to obtain a return on the investment. These resources may facilitate auditors' knowledge and skill acquisition process and help them to develop industry-specific expertise (Francis 2011). Therefore, audit firms that invest in the industries in which they concentrate their services are argued to develop significant industry-specific expertise. These industry specialist audit firms are subsequently more likely

[38] I use both continuous and dichotomous measures, which I detail and describe the measurement of in Chapter 5.

to exhibit improved judgment and decision making when assessing clients' inherent risks, conducting analytical procedures and obtaining and evaluating evidence (Dowling and Leech 2007; O'Donnell and Schultz 2003; Janvrin et al. 2008), which may lead to superior audit outcomes. Therefore, the portfolio-share measure of industry specialisation is argued to capture the provision of higher audit quality (Neal and Riley 2004; Numan and Willekens 2012). In practice, large audit firms have recognised the importance of industry expertise and concentrated their services in particular industries (KPMG 2014, PWC 2014); thus, it is also arguable that the portfolio-share measure most directly reflects such business decisions.[39]

As with other measures, the portfolio-share measure has potential limitations. First, the measure is mathematically affected by the relative size of the industry (Neal and Riley 2004). Under traditional applications of this method, an audit firm is more likely to be identified as a specialist for clients in industries where the total industry fee base is large relative to other industries in the audit firm's potential portfolio.[40] Consequently, I control for this limitation of the portfolio-share method using two-stage regression analyses, which account for the relative size of the fee base in the client's industry. Another potential weakness of the portfolio-share measure is that the effect of knowledge spillovers might not be captured as cleanly as under the

[39] Further, client's size relative to the rest of the industry systematically affects the identification of specialists under the market-share measure, but does not impose similar effects on the portfolio-share measure. The portfolio-share measure may potentially recognise relatively small audit firms as specialists in industries where they may generate most revenue, even if they do not have leading market share in that industry (Neal and Riley 2004). However, since I restrict my sample to clients audited by a Big N firm, this advantage of the portfolio-share measure over the market-share measure is not of direct relevance to my thesis.

[40] Assume, for example, that an auditor serves three clients: Firm 1 is a member of industry A, and Firms 2 and 3 are members of industry B. Industry A is an industry comprised of many large firms that typically pay very high audit fees (e.g. Firm 1 pays the auditor \$50 million). Industry B is a small industry in which Firm 2 and Firm 3 pay \$5 million and \$10 million in audit fees, respectively, to that auditor. According to the portfolio-share metric, the auditor is recognised as having specialisation in Industry A (76.92%: 50 m / (50 m+5+10), relative to Industry B (23.08%: 15 m / 65 m), although the auditor services two clients in Industry B.

market-share measure, as the value of spillovers may be more directly a function of the aggregate number of (fee revenue from) similar clients, rather than the proportion of similar clients in the auditor's portfolio.

In summary, I maintain that the portfolio-share measure of audit firm industry specialisation most closely reflects the economic construct of specialisation (as proposed by Smith 1776). It also captures the audit firm's expertise in industries in which the firm makes a greater relative investment. This is consistent with the industry-focused objectives asserted by the large audit firms that are the focus of my study. Further, while the market-share measure of industry specialisation is increasing with market dominance, which does not necessarily lead to positive audit outcomes, such a problem is not of direct relevance when the portfolio-share measure of industry specialisation is used. Empirical auditing research frequently uses the portfolio specialisation measure to proxy audit quality and finds a positive relation between this measure and desired audit outcomes (Krishnan 2001, 2003; Peters et al. 2001; Abbott et al. 2005; BCK 2008; Payne 2008; Numan and Willekens 2012). Therefore, my thesis uses the portfolio-share measure as the primary proxy for audit firm industry specialisation, and controls for its limitations accordingly.

4.2.3 Refinements of the Market-share and Portfolio-share Measures

Recent auditing research has proposed refinements of the basic market-share and portfolio-share measures, as described below.

4.2.3.1 Weighted market-share measure

Owing to the potential problems associated with the exclusive use of either the market-share or the portfolio-share measure, Neal and Riley (2004) propose an alternative measure that captures the complementary relation between the two measures. Noting the inconsistent results reported in prior research across these proxies, Neal and Riley (2004) argue that these two measures may capture different aspects of auditor expertise, and may act as complements rather than substitutes. Thus, they propose a 'weighted market-share' measure equal to the product of the audit firm market-share and portfolio-share measures:

Weighted Market share = Market share * Portfolio Share

Under this approach, the previously identified shortcomings of the individual measures can be mitigated. For example, if an audit firm has a high portfolio share with respect to firms in a large industry, this may simply reflect the size of the industry (and the fees payable within it). The weighted market-share measure will only be 'high' if the audit firm also has high market share in that industry. Likewise, if an audit firm has a very high market share in some industries, this may possibly reflect the level of audit market concentration or the size of the client firms in a particular industry. The weighted market share again will only be 'high' if the audit firm has high portfolio share in that industry. Therefore, the weighted market-share approach captures some of the attributes of both approaches. This approach is later used in an examination of audit firm industry specialisation and accounting restatements by Romanus et al. (2008), who find that industry specialisation is negatively correlated with the likelihood of accounting restatements.

Neal and Riley's (2004) complementary measure is based on the assumption that both market share and portfolio share are increasing in industry expertise and are thus positively related to audit outcomes. Therefore, if one of the measures does not capture audit quality, and thus does not lead to positive audit outcomes, the weighted market-share approach will reduce the explanatory power of the single specialisation measure that is appropriately defined. As I have already argued that greater market share may be associated with lower audit quality in some cases, this limitation may also affect the complementary measure similarly. Thus, I use the weighted market share as a supplementary, rather than primary, measure of audit firm industry specialisation.

4.2.3.2 City-based industry specialisation measures

Both the market-share and portfolio-share measures described above have traditionally been estimated across national audit markets. As such, each implicitly assumes that the industry expertise of accounting firms is transferrable across different offices within audit firms. However, more recently, academic research argues that audit firm's industry expertise is largely city-specific and is not easily transferable to other offices throughout the firm's network (Ferguson et al. 2003; Francis et al. 2005b; Reichelt and Wang 2010). The above argument is based on the following contentions. First, deep industry expertise resides in the unique individual audit partners and staff, and is thus limited to their offices of practice, which most often service clients predominantly in one locale (Ferguson et al. 2003). Second, in the U.S. where the audit market is quite decentralised and Big N audit offices are widespread across the country, it may be difficult to maintain a uniform firm-wide expertise across all offices (Francis et al. 2005b). Thus, deep industry expertise is argued to be city-specific.

Empirical research provides some evidence in favour of the superiority of city-specific expertise measures. Using the audit firm's fee premium to explore national and city-level industry expertise, Ferguson et al. (2003) find a fee premium only for Australian audit firms that are both national and city market-share leaders. A similar study based on U.S. data reports a fee premium for city-specific industry specialists who are not national industry specialists (Francis et al. 2005b).[41] In addition, Reichelt and Wang (2010) find lower discretionary accruals for clients audited by city-specific industry specialists relative to national industry specialists, and even lower discretionary accruals in clients audited by joint national and city-specific industry specialists. Digging deeper, Francis and Yu (2009) claim that auditor's expertise is dependent on their office size. Audit firms with larger offices and greater engagement hours provide their audit staff with superior local support and more opportunities to consult with peers, which facilitates the development of 'in-house' expertise and results in better audit outcomes. Their empirical results show that Big 4 audit firms' office sizes are positively associated with audit outcomes measured by the likelihood of issuing a going-concern opinion and lower levels of absolute discretionary accruals (Francis and Yu 2009). Since city-level audit firm data are

[41] Francis et al. (2005b) explain that the inconsistency in their findings relative to Ferguson et al. (2003) is mainly driven by the institutional differences between Australia and the U.S. and the decentralisation of the U.S. audit market. Specifically, Australian accounting standards and reporting requirements are less detailed, and the regulatory body, the Australian Securities and Investments Commission, is relatively less aggressive in regulating companies than is the Securities and Exchange Commission in the U.S. In addition, while most Australian publicly listed companies and Big N audit firms have headquarters in Sydney, Melbourne and Perth, clients and audit firms are more geographically dispersed in the U.S.

only available from Audit Analytics after 1999,[42] my main analysis focuses on national-level data, although I conduct city-level tests in my sensitivity analysis.

4.3 Analyst Quality

Analyst quality is conceived in my thesis as reflecting analysts' access to private information[43] and analysts' superior ability to identify and process relevant complex information. While analyst quality cannot be observed directly, prior studies identify numerous proxies, including analysts' general and firm-specific forecasting experience, the size of the analyst's employer (brokerage size), analyst prior forecast accuracy, analysts' portfolio complexity and measures based on rankings published in the financial press (the 'All-Star' status). I now describe these proxies in detail and analyse the extent to which each proxy may capture underlying analyst quality.

4.3.1 Analyst Experience

Experience is 'the knowledge or skill acquired by a period of practice of something especially that gained in a particular profession' (*Oxford English Dictionary* 2010, 615). As knowledge and skill have positive effects on individuals' decision making and performance, experience is argued to improve judgment and decision making quality (Bonner 2008). Empirical research documents a positive relationship between

---

[42] A city is defined as the U.S. Census Bureau's definition of metropolitan statistical areas (MSA). Due to data limitations, Francis et al. (2005b) assume that the engagement office of an audit firm is located in the same city as the clients' headquarters and they thus simply use the city codes of clients' headquarters as reported in *COMPUSTAT* to represent those for audit firms, and match these codes with MSA codes. A more precise measure of auditor city developed in recent studies (i.e. Francis and Yu 2009; Reichelt and Wang 2010) identifies auditors' cities from Audit Analytics and matches them with MSA codes. My thesis uses this more precise and recent method to measure city-specific auditor industry specialisation. More details are provided in the sensitivity analysis section in Chapter 7.

[43] As footnoted in Chapter 3, analysts' access to private information sourced from within the covered firms was specifically proscribed by the Regulation Fair Disclosure (Reg FD) on 23 October 2000. However, private information need not refer only to that sourced from within the covered firms. Superior analysts may have superior contacts in other areas, such as in industry advocacy bodies, economic consultancies and the business press firms.

experience and performance in several disciplines, including clinical psychology (Oskamp 1965), physics (Chi et al. 1982) and auditing (Frederick and Libby 1986; Hamilton and Wright 1982).

In the analyst literature, experience is measured in terms of the period over which an analyst has issued forecasts, either for any firm (general experience) or for a specific firm (firm-specific experience). Analysts who acquire greater knowledge and skills over time are argued to be more capable of identifying and processing complex information and making accurate forecasts (Jacob et al. 1999; Clement and Tse 2003, 2005; Drake and Myers 2011). Further, analysts with greater experience may be better able to develop strong relationships with the covered firms and (prior to the enactment of Regulation Fair Disclosure [Reg FD] at least) obtain more relevant information about the covered firms outside those reported on the financial reports (Clement 1999). The following sections describe these experience measures and review the literature relevant to their interpretation.

4.3.1.1 General experience

General experience refers to the period over which an analyst has issued forecasts for any firm and is argued to increase analyst quality and improve analyst forecast accuracy (Clement 1999, Clement and Tse 2003; 2005; Casey 2012; Kim et al. 2011; Brown and Mohammad 2010; Drake and Myers 2011). Clement (1999) documents a significant positive relation between analysts' general experience and individual analysts' absolute forecast errors.[44] Kim et al. (2011) argue that the documented

[44] Clement (1999) uses a relative measure of analysts' absolute forecast error, defined as the difference between the analyst's absolute forecast error for a given firm and the mean absolute forecast errors for that firm, divided by the latter.

relationship between analyst quality proxies and forecast accuracy may also be attributable to the frequency with which analysts update their forecasts late in the reporting year. They find that analysts' general experience increases the likelihood that the analyst will revise forecasts later in the quarter and thus reduces analysts' absolute forecast errors.[45]

Analyst general experience is also found to be positively related to analysts' other forecasting performance. Clement and Tse (2005) investigate the relation between analyst quality and forecast boldness. They identify 'bold forecasts' as those deriving from forecast revisions that result in the new forecast being more extreme than both the analyst's own previous forecasts and extant consensus forecasts. Thus, bold forecasts reflect cases in which the analyst's opinion diverges from consensus. Conversely, 'herding forecasts' are those that are issued at a level between the analyst's prior forecasts and the consensus forecasts. Clement and Tse's (2005) results show that analysts with greater general experience are less likely to herd, and are more accurate. Further, Drake and Myers (2011) refine the earlier study of Bradshaw et al. (2001) to investigate the impact of analyst quality on the relation between forecast optimism and the accrual component of earnings. Bradshaw et al. (2001) find that analysts are misled by the mean reversion implications of large accruals, or that they deliberately collaborate with management to inflate forecasts. It is suggested that the market inefficiency in pricing the components of earnings (Sloan 1996) may be attributable to analysts' behaviour. Drake and Myers (2011) provide evidence that more experienced analysts exhibit lower accrual-related long-

[45] Kim et al.'s (2011) measure of analysts' absolute forecast errors is the numerator in Clement's metric.

horizon forecasts optimism, suggesting that these analysts predict the resolution of accruals more accurately than do others.[46]

4.3.1.2 Firm-specific experience

Firm-specific experience refers to the period over which an analyst has issued forecasts for a specific firm and is argued to increase analysts' understanding of the idiosyncrasies of covered firms' financial performance and reporting practices, thereby improving forecast accuracy. However, the reported results with regard to the relation between firm-specific experience and forecasting performance are inconsistent (Mikhail et al. 1997; Clement 1999; Keskek et al. 2013; Jacob et al. 1999). These prior results are detailed below.

Some prior studies posit and document evidence that analysts' firm-specific experience improves forecast accuracy, while others show that the impact of firm-specific experience on forecast accuracy is not significant. For example, Mikhail et al. (1997) find that short-horizon forecast accuracy is positively correlated with analysts' firm-specific experience, and Clement (1999) reports a positive relation between forecast accuracy and analysts' firm-specific experience. However, examining analyst forecast accuracy over a sample period similar to Clement's (1999), Jacob et al. (1999) find no evidence that forecast accuracy increases with analysts' firm-specific experience. This difference in results is likely due to differences in these studies' research designs. For example, while Clement (1999) controls for firm-year fixed effects, Jacob et al. (1999) control for analyst aptitude

[46] Drake and Myers (2011) find that their main results are consistent with either an expertise/ability effect or an information effect, but they produce additional evidence (pre-post-Reg FD tests) suggesting that the information effect is the most likely driver of their results.

and analyst–broker and analyst–company alignment by including analyst fixed effects, brokerage firm fixed effects and analyst–company and analyst–industry fixed effects.[47] Keskek et al. (2013) show that analysts' firm-specific experience is correlated with analyst forecast accuracy in the pre-Reg FD period, but not in the post-Reg FD period. They interpret this result as evidence that analysts' (former) access to management's information drives the firm-specific experience effect. Therefore, as an additional analysis, I employ the Reg FD reforms as an exogenous shock to the importance of my analyst 'quality' measures, and re-test my hypotheses accordingly.

Examining other analysts' forecasting properties, prior studies find no evidence that firm-specific experience affects performance. Clement and Tse (2005) demonstrate that analysts with greater general experience are less likely to herd, but find no association between analysts' firm-specific experience and forecast boldness (as against herding forecasts). Further, Drake and Myers (2011) find no evidence that analysts with greater firm-specific experience exhibit lower accrual-related forecast optimism.

In conclusion, the impact of analysts' firm-specific experience on forecasting performance observed in earlier studies is likely to reflect differences in research design and sample period. Although the empirical evidence regarding the relationship between firm-specific experience and forecasting performance is mixed,

[47] My thesis examines the association between audit quality, forecast accuracy and analyst quality on a firm-year or analyst-firm-year basis. I either calculate the average values of the analyst quality proxies at the firm-year level, or use the values of quality proxies to identify the 'worst' and 'best' quality analysts following a given firm. Therefore, Jacob et al.'s (1999) control variables, which are particularly relevant to a study examining individual analysts' forecasting performance, are of little relevance to my thesis.

the theoretical argument underpinning a positive association between analysts' firm-specific experience and analysts' forecasting performance appears sound. For this reason, I include analysts' firm-specific experience as a proxy for analyst quality.

4.3.2 Brokerage Size

The size of the firm for which an analyst works is another factor posited to affect the properties of analysts' forecasts. Large brokerage firms are able to provide the analyst better support, including administrative assistance, training opportunities, superior datasets, evaluation procedures and techniques to analyse data (Drake and Myers 2011; Clement 1999). In addition, large brokerage firms may have developed strong relationships with covered firms over time and possess a greater amount of relevant information about a firm (Clement 1999; Jacob et al. 1999). Thus, analysts working for these larger brokerage firms may have (or have previously enjoyed) more information about a covered firm and superior skills and support relevant to their forecasting decisions. These analysts are argued to generate forecasts that are more accurate (Jacob et al.1999; Clement 1999).

Brokerage firm size has been shown to affect analysts' forecasting performance positively. Jacob et al. (1999) and Clement (1999) each reports a positive association between brokerage size and analyst forecast accuracy. Examining the impact of brokerage size on other forecasting properties, prior studies show that analysts from larger brokerage firms are more likely to make bold forecasts (Clement and Tse 2005) and issue more timely forecasts (Kim et al. 2011).

However, other recent studies find either a weak or an insignificant relationship between brokerage firm size and analyst forecasting properties. Drake and Myers (2011) show that accrual-related over-optimism decreases with brokerage size, although the significance of the negative association between brokerage size and forecast over-optimism disappears after controlling for the number of analysts following a firm as well as firm size and firm age. Keskek et al. (2013) find no relationship between brokerage firm size and forecast accuracy in the post-Reg FD periods. They further demonstrate that brokerage size in their sample is associated with a decrease in analysts' forecasting abilities (as measured by prior forecast accuracy and forecast boldness) and an increase in analysts' portfolio complexity (measured by the number of firms followed by an analyst) in the post-Reg FD periods. This evidence reinforces the importance of examining the impact of Reg FD in my sample as a sensitivity analysis.

4.3.3 'All-Star' Status

Large buy-side investment institutions have an incentive to identify outstanding analysts who provide more accurate forecasts and recommendations, and they consequently subscribe to services that rank analysts. Two analyst-ranking systems are particularly well known: the All-American Research Team rankings published by *Institutional Investor* (*II*), and the 'Best on the Street' list published by *The Wall Street Journal* (*WSJ*). Analysts selected to appear in these rankings/lists are colloquially described as 'All-Star' analysts. These analysts are considered to have superior ability and/or resources, which contribute to their better forecasting performance. The ranking procedures are different in the two systems and described separately below.

Under the *II* ranking procedures, surveys are distributed each year in March, April or May to fund managers and directors of research in U.S., European and Asian investment funds. Survey respondents are required to submit votes for the four most helpful or valuable analysts by recalling their names (i.e. they must write the analysts' names on the form). *II* requires respondents to evaluate analysts based on a combination of 'hard' attributes (earnings forecasts and stock selections) and soft attributes (e.g. industry knowledge, special services, and quality of sales force). In the October 2010 issue of *II's* magazine, *II* explained that their:

> Rankings were determined strictly by using numerical scores. We took the number of votes awarded to each analyst and weighted the votes based on the size of the institution responding and the place it awarded to that analyst (first, second, third or fourth) … We consulted nearly 3,500 individuals at some 970 firms, including more than 90 of the 100 biggest U.S. equity managers. Our respondents manage an estimated $10.2 trillion in U.S. equities.

By contrast, the *WSJ* ranking system ranks analysts according to a single criterion: the industry-adjusted returns on portfolios constructed from the analyst's recommendations during the year and prior to the ranking. The *WSJ* ranks the top five analysts in each industry subject to the constraint that analysts cover at least five firms in that industry. Owing to these differences in the ranking procedures, *WSJ* typically recognises more analysts from smaller brokerage firms as 'All-Star' (Bagnoli et al. 2005).[48] However, for the purposes of my thesis, which examines the factors affecting earnings forecast accuracy (rather than investment

[48] The *II* ranking procedures ask survey respondents to write down analysts' names. Analysts from larger brokerage firms are more likely to be known either through the brokerage firms' relationships with large fund managers or through these firms' advertising force. Therefore, this ranking system favours analysts from large brokerage houses.

recommendations), the *II* ranking system is of most direct relevance; it is thus the only ranking system that I employ.

A large body of academic research explores the performance of 'All-Star' analysts on forecasting performance. Stickel (1992) obtains forecasts from Zacks database and find that *II* 'All-Star' analysts generally forecast more often and more accurately than do other analysts in the selection year. Chan et al. (2004) study the differences in forecast accuracy over a three-year period between 'All-Star' analysts, selected from either *II* or *WSJ*, and their counterparts for the period 1992–2001. They show that 'All-Star' analysts consistently provide more accurate short-horizon forecasts than do others in the year before the 'All-Star' selection, in the selection year and in the post-'All-Star' selection year.

4.3.4 Portfolio Complexity

Analysts' portfolio complexity increases with the number of firms and industries followed by an analyst (Clement 1999; Clement and Tse 2003, 2005; Keskek et al. 2013; Drake and Myers 2011). There are competing arguments regarding the impact of analysts' portfolio complexity on forecast accuracy. Studies proposing a negative relationship between analysts' portfolio complexity and forecasting performance consider that analysts' time and resources are scarce and, as such, that following fewer firms and industries allows more time and attention to be devoted to each covered firm, helping these analysts to develop close relationships with firms and acquire a greater amount of relevant information (Clement 1999; Drake and Myers 2011). Therefore, analysts' portfolio complexity is typically viewed as an inverse proxy for analyst quality, and thus as negatively related to forecast accuracy.

Conversely, considering the capability, specialisation and diversification effects, the empirical association between analysts' portfolio complexity and forecasting performance may be unclear or positive. Prior studies argue that more capable (and thus higher quality) analysts are likely to be assigned greater responsibilities, as reflected in the analyst's portfolio (Jacob et al. 1999). For example, Stickel (1992) shows that *II* 'All-Star' analysts follow a median of 14 firms, while non-star analysts follow a median of only eight firms.[49] Although Clement (1999) and Drake and Myers (2011) make opposite predictions about portfolio complexity and forecasting performance, they show that more experienced analysts (who are likely of higher quality) cover more firms and industries. As discussed above, these higher-quality analysts (more experienced, or awarded 'All-Star' status) possess superior ability to identify and process relevant information. As such, it is unclear whether the capability effects outweigh the information disadvantage associated with portfolio complexity.

In addition to the capability argument, analysts following multiple firms within the same industry are argued to benefit significantly from knowledge spillovers across the covered firms and to develop specialisation in that industry (Jacob et al. 1999; Siegel et al. 2011; Gilson et al. 2001). Gilson et al. (2001) explicitly regard an analyst as a specialist with respect to a covered firm if that analyst follows at least five other firms within the same industry as the sample firm. Further, Kini et al. (2009) argue that there is a trade-off between the relative information benefits of simple and complex portfolios. They propose that analysts following multiple

[49] In my sample, *II* 'All-Star' analysts follow more firms than do non-star analysts, with the difference being statistically significant. I also find that more experienced analysts cover more firms and industries than do inexperienced analysts.

industries (complex portfolio) may reduce some of the information benefits associated with following a simple portfolio, but gain exposure to complementary information and acquire knowledge from the coverage of related industries. If the benefits from the complementary information and knowledge are relatively important, analysts with more complex portfolios may forecast more accurately. Thus, 'busy' analysts may actually have a greater amount of relevant information and the superior ability to process such information, increasing the accuracy of their issued forecasts. The empirical results are discussed.

Some empirical evidence shows that portfolio complexity is negatively associated with forecasting performance. Clement (1999) documents evidence that analyst forecast accuracy decreases with analysts' industry and firm coverage. Similarly, Jacob et al. (1999) report a negative association between forecast accuracy and analysts' firm coverage, although they propose an alternative argument as stated above. According to Kim et al. (2013), the number of firms and industries an analyst follows is negatively associated with more timely forecasts (i.e. forecasts tend to be revised later in the forecasting period) and forecast accuracy.

Other studies find weak or no evidence that analysts' portfolio complexity is associated with forecasting performance. Siegel et al. (2011) provide evidence that portfolio complexity (measured by the number of industries an analyst follows) is not associated with analyst forecast accuracy after controlling for analyst specialisation in business segments (i.e. the number of firms' business segments in the industry followed by an analyst relative to the total number of segments for all firms followed by that analyst). Clement et al. (2003) examine the relationship between analyst quality and forecast accuracy across 10 countries, and find that the number of firms followed by an analyst positively affects forecast accuracy in only three of the non-U.S. countries. Using a two-stage Heckman approach to account for the fact that portfolio choice is endogenously determined, Kini et al. (2009) show that analyst forecast accuracy increases when analysts diversify their coverage portfolio across industry sectors (more complex portfolio). Examining other analyst forecast properties, Clement and Tse (2005) document evidence that analysts who follow a large number of industries are less likely to make bold forecasts; these authors show no evidence that the number of firms followed by an analyst is associated with forecast boldness. Drake and Myers (2011) demonstrate that after controlling for firm characteristics, analysts who follow fewer firms have a better understanding of the accruals towards which they show less over-optimism. However, this relationship does not hold when industry coverage is used to indicate analysts' portfolio complexity.

Since prior studies report competing arguments and predictions regarding the impact of analysts' portfolio complexity on forecast accuracy and demonstrate inconsistent evidence, I do not use this measure as a proxy for analyst quality. However, as portfolio complexity may be correlated with forecast accuracy and other analyst quality proxies discussed above, I include this in my empirical tests. For example, analysts employed by a larger broker may follow a smaller number of firms and industries and those ranked as 'All-Star' may follow a greater number of firms, portfolio complexity variables thus could be correlated with brokerage size and 'All-Star' status. Experience proxies may be correlated with portfolio complexity variables if more experienced analysts follow more complex portfolios (Clement 1999).

4.3.5 Analyst Prior Forecast Accuracy

Analyst prior period forecast accuracy is also frequently used in the literature as a proxy for analyst quality. Evidence suggests that analysts with superior prior period forecast accuracy are likely to issue forecasts in the current period that are more accurate, more bold (Clement and Tse 2005; Keskek et al. 2013) and more timely (Kim et al. 2013). However, analyst prior forecast accuracy is not appropriate for use in my thesis, as it represents the lagged value of my dependent variable. As my main test variable (audit quality) is relatively 'sticky' across time, if analyst prior accuracy were to be included in the regression, this would likely cause an endogeneity problem, as any factors (such as audit quality) that explain current year forecast accuracy might also be the cause of analyst prior accuracy. For this reason, analyst prior accuracy is not used to proxy analyst quality in my study.

4.3.6 Analyst Proxies Used in My Thesis

My thesis uses analysts' general and firm-specific experience, brokerage size and 'All-Star' status as alternate proxies for analyst quality. As discussed earlier in this chapter, each of these proxies reflects the properties underlying the construct for 'analyst quality', which advances forecast accuracy. I also develop composite score measures to capture the attributes of an analyst along the four identified dimensions to assess the analyst's overall quality. The precise measurement of each of these variables is detailed in the following chapter.

4.4 Chapter Summary

This chapter introduced and described the proxies for audit firm industry specialisation and analyst quality used in my thesis. Section 4.2 identified and discussed the various measures of audit firm industry specialisation (as a proxy for audit quality) and emphasised the superiority of the portfolio-share measure over the other measures as to capturing audit firm industry expertise and thus superior quality audit services. Section 4.3 described several measures of analyst quality identified in the literature and argued that general experience, firm-specific experience, brokerage size and *II* 'All-Star' status are the measures of most direct relevance to my thesis. In the following chapter, I introduce and discuss the regression models and estimation methods used to test my hypotheses. I also explain in detail the measurement of my dependent, test and control variable.

CHAPTER 5: RESEARCH DESIGN

5.1 Introduction

In Chapter 3, I developed my hypotheses regarding the impact and importance of audit firm industry specialisation on overall analyst forecast accuracy and the extent to which the relationship between audit firm industry specialisation and forecast accuracy varies cross-sectionally with the underlying riskiness of the client firm's operations and the quality of the analysts covering the client firm. The empirical proxies for auditor industry specialisation and analyst quality were described in Chapter 4. In this chapter, I describe my research design. First, Section 5.2 provides an overview of the general form of the regression models employed to test my hypotheses and the structure of the data used. I then describe the measurement of variables in Section 5.3. Finally, in Sections 5.4 and 5.5, respectively, I discuss the details of the estimation methods and the specific regression models employed. Section 5.6 concludes the chapter.

5.2 Overview of the Models and Methods

This section describes the general form of the regression models used to test my hypotheses. These simpler formulations of my regression models are provided to contextualise the following descriptions of variable measurement (Section 5.3) and the choice of regression method (Section 5.4). The complete (detailed) specifications of the various regression models are provided in Section 5.5.

Hypotheses 1a and 1b make predictions about the relationship between audit firm industry specialisation and consensus analyst forecast accuracy. H1a is non-directional due to the competing effects of audit quality on the reliability of financial statements as an information source and on the client's ability to manipulate current earnings. However, H1b conjectures a positive association between analyst long-horizon forecast accuracy and audit firm industry specialisation, as benchmark-beating incentives are much weaker with respect to forecasts redundant at the end of the financial year. While H1a and H1b differ in the forecast horizon examined and with respect to the time at which industry specialisation is measured, the composition of the empirical models used to test these hypotheses is otherwise similar. I regress analysts' absolute short- or long-horizon forecast errors on proxies for audit firm industry specialisation and a set of controls previously found to be associated with forecast errors.[50] I use both continuous and dichotomous measures of audit firm industry specialisation. The general form of these models is detailed in Equation (1):

$$ABSFE = \beta_0 + \beta_1 INDSP + CONTROLS + \varepsilon \qquad (1)$$

where

ABSFE	=	analysts' absolute forecast errors, measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price, measured either immediately before current earnings are announced (H1a) or immediately after prior period earnings are announced (H1b); and
INDSP	=	audit firm industry specialisation, measured as the sum of the square root of the total assets of the clients that an audit firm services in a particular industry, relative to the sum of the square root of the total assets of all clients of that audit firm.

[50] Analysts' absolute forecast errors are an inverse measure of analyst forecast accuracy and are frequently used in the literature (Payne 2008; Hall and Tacon 2010; Dhaliwal et al. 2012). The details of the control variables are discussed in Section 5.3.6.

To support H1a, which predicts that audit firm industry specialisation *(INDSP)* is associated with short-horizon forecast errors, *INDSP* is expected to have a significant coefficient (β_1) of either sign. A significant negative coefficient for *INDSP* is predicted under H1b, which argues that analysts' long-horizon forecasts are more accurate (absolute forecast errors are lower) when audit quality is higher.

Hypothesis 2 predicts that audit firm industry specialisation has a greater effect on forecast errors when the underlying uncertainty surrounding a client firm's business transactions is greater, because in these circumstances a high-quality auditor has a greater potential impact on accrual estimation errors and the overall quality of clients' supporting disclosures. While analysts' absolute forecast errors are expected to be decreasing in audit quality, this negative relationship is predicted to be stronger when client's operating risk is higher. To test this hypothesis, proxies for clients' operating risk and their interaction with audit firm industry specialisation are added to the model introduced above, as per Equation (2):

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 CLIENT\ OPERATING\ RISK + \beta_3 INDSP*CLIENT\ OPERATING\ RISK + CONTROLS + \varepsilon \quad (2)$$

If the negative association between audit firm industry specialisation and analysts' absolute forecast errors is stronger when client's operating risk is high as per H2, the coefficient (β_3) for the interaction term is expected to be significantly negative.

Hypotheses 3a and 3b predict that industry specialist auditors affect the relative forecast accuracy of high-quality and low-quality analysts. H3a applies this general proposition to the differential impact of audit firm industry specialisation on forecast accuracy across firm-years for which the *average* quality of the cohort of analysts following a firm differs. I test H3a by expanding the model used to test H1b to include proxies for the mean quality of analysts following a client firm in a particular year, and the interactions between these proxies and audit quality. The various proxies for analyst quality and their interactions are included in a single regression model (Equation [3]) to control for the potential collinearity among these proxies (Clement 1999).

$$ABSFE = \beta_0 + \beta_1 INDSP + \Sigma ANALYST\ QUALITY + INDSP * \Sigma ANALYST\ QUALITY + CONTROLS + \varepsilon \quad (3)$$

Where I use a composite measure derived from the specific analyst quality proxies, this composite measure and its interaction with audit firm industry specialisation are included in Equation (4).

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 ANALYST\ QUALITY + \beta_3 INDSP * ANALYST\ QUALITY + CONTROLS + \varepsilon \quad (4)$$

Hypothesis 3a is supported if the coefficient(s) for the interaction(s) between audit firm industry specialisation and analyst quality proxy is (are) positive and significant, indicating that audit quality has a smaller (greater) impact in improving forecast accuracy when the average quality of analysts is higher (lower).

Hypothesis 3b proposes that greater audit firm industry specialisation reduces the difference in forecast accuracy between the 'worst' quality analyst and the 'best' quality analyst following a client firm in a given year. I argue that a high-quality audit should have a greater impact on the accuracy of the 'worst' analyst, who may depend more heavily on the quality of the published audited financial statements when making forecasts. Equation (5) regresses the difference in forecast errors between the 'worst' quality and 'best' quality analyst (estimated separately for each of the individual analyst quality proxies) following a client firm in a given year on audit firm industry specialisation and control variables.

$$DIFABSFE = \beta_0 + \beta_1 INDSP + CONTROLS + \varepsilon \qquad (5)$$

Where

DIFABSFE = the absolute forecast error of the 'worst' quality analyst minus the absolute forecast error of the 'best' quality analyst where the 'worst' and 'best' quality analysts are determined with respect to a particular analyst quality proxy.

A significant negative coefficient for *INDSP* would support H3b, indicating that higher audit quality increases the usefulness of the current year's earnings for predicting future earnings, in turn reducing the difference in forecast accuracy across the 'worst' and 'best' quality analysts.

The models described above are estimated by applying regression techniques to archival data. The composition of the models and the structure of the sample suggest two main concerns regarding the choice of regression method. First, the sample consists of repeated observations of the same firms across several years. A simple pooling of the sample may result in inconsistent estimates due to unobserved firm or time effects. A firm effect occurs when the residuals for a given firm are correlated across years, while a time effect occurs when the residuals for a particular year are correlated across different firms (Wooldridge 2002). Ordinary Least Squares (OLS) estimators are unbiased only if the residuals generated are independent and identically distributed. If this is not the case, adjustments must be made by either transforming the data to remove the dependence (e.g. the use of firm and/or time fixed effects) or adjusting the estimated standard errors to reflect within-subject and/or within-year clustering (Peterson 2009, 435). My study controls for firm effects by adjusting OLS standard errors for within-firm clustering. Time effects are controlled either by adjustment of standard errors for clustering or by the inclusion of time fixed effects depending on the nature of the data used to test a particular hypothesis.

The second issue of concern is the possible endogenous determination of the key test variable, audit firm industry specialisation. Godfrey and Hamilton (2005) show that the selection of an industry specialist auditor is not random; it is affected by several factors including the client firm's research and development expenditures, length of operating cycle, size, leverage and earnings per share. The factors affecting the choice of an industry specialist auditor may be correlated with the factors affecting the accuracy of analysts' forecasts and their effect may not be fully captured in the vector of control variables. If this is the case, an endogenous selection bias may exist (i.e. the factors affecting the selection of an industry specialist auditor may be correlated with the residuals from the regressions described above). Consequently, if standard OLS regression were applied to these models, this might generate inconsistent estimators. To address this potential endogeneity threat, I employ two-stage regression approaches, including two-stage least squares (2SLS), propensity score matching (PSM) and Heckman treatment-effect regressions. These methods are described in detail in Section 5.4.

5.3 Measurement Issues and Data Sources

In the following sub-sections, I describe the measurement of the variables employed in my regression models. The measurement of the dependent variables, analyst forecast accuracy and the difference in forecast accuracy of the 'worst' and 'best' analysts, are first described (Section 5.3.1), followed by the proxies for audit quality (Section 5.3.2). An illustration of the temporal relation between these two variables is provided in Section 5.3.3. Sections 5.3.4 and 5.3.5 describe the measurement of the variables used to test Hypotheses 2, 3a and 3b; namely, firm operating risk and analyst quality. Finally, the control variables and their measurement are identified and described in Section 5.3.6.

5.3.1 Analyst Forecast Accuracy

The majority of my empirical tests focus on explaining the properties of analysts' consensus long-horizon forecast errors. Long-horizon consensus forecasts are measured by the mean value of each individual analyst's first one-year-ahead EPS forecast made subsequent to the prior year's earnings announcement date. To reduce the noise caused by other information, forecasts that are issued more than 90 days after the prior year's earnings reporting date are excluded. Short-horizon consensus forecasts, used to test H1a only, are measured by the mean of the final forecasts made by each individual analyst in the 90 days immediately prior to the earnings reporting date. Analysts' forecasts and their respective actual values are obtained from the *I/B/E/S* Detail History file. Following prior literature (Payne 2008; Hall and Tacon 2010; Dhaliwal et al. 2012), forecast accuracy is measured as the absolute value of the forecast error, where the forecast error is the difference between the mean analyst forecast EPS and the actual EPS as reported by *I/B/E/S* (this is an inverse function of analyst forecast accuracy, as a larger value indicates a less accurate forecast).[51] Since prior literature differs concerning the scaling of forecast accuracy, I use two measures: an undeflated measure and a measure of forecast accuracy deflated by stock price, as per Equations 6a and 6b below:

$$ABSFE\_un = |Actual\ EPS - Forecast\ EPS| \quad (6a)$$

$$ABSFE = |(Actual\ EPS - Forecast\ EPS) / P| \quad (6b)$$

Where

ABSFE_un	=	undeflated absolute forecast errors;
ABSFE	=	price-deflated absolute forecast errors;
Actual EPS	=	reported EPS from *I/B/E/S* Detail File;
Forecast EPS	=	mean of outstanding analyst forecasts for annual EPS ; and
P	=	stock price one month prior to reporting date

My tests of H3b employ the difference in forecast errors between the 'worst' and 'best' quality analysts (*DIFABSFE*) as the dependent variable, where the 'worst' and 'best' analysts are identified according to each individual analyst quality proxy, as described in Chapter 4. I provide an example of the detailed calculations of *DIFABSFE* in Section 5.3.5.

DIFABSFE = the absolute forecast error of the 'worst' quality analyst (*ABSFE_W*) minus the absolute forecast error of the 'best' quality analyst (*ABSFE_B*), where the 'worst' and 'best' analysts are determined based on the values of each analyst quality proxy. (7)

I conduct a separate test of H3b for each analyst quality proxy.

[51] For consistency, both the forecast earnings and actual earnings for calculating forecast errors were collected from *I/B/E/S*, each of which is based on the opinion of the majority of analysts regarding the inclusion/exclusion of non-recurring items from the earnings measure, and which often differ from GAAP earnings (Bradshaw and Sloan 2002). In addition, the *I/B/E/S* data are adjusted to reflect the effects of stock splits and other dilution factors.

5.3.2 Proxies for Audit Firm Industry Specialisation

Industry specialists are argued to apply valuable resources and industry-specific knowledge and skills to audit processes, thereby improving audit quality as evidenced by outcomes such as financial reporting quality. I use the auditor portfolio share as my primary measure of industry specialisation for reasons detailed in Chapter 4,[52] and also because this measure is common to the most directly relevant prior papers (Payne 2008; BCK 2008). The portfolio-share industry specialisation measure attempts to capture the extent to which audit firms concentrate their productive activities (and thus fee base) within particular industries, thus capturing auditor expertise in industries within which auditors are demonstrably willing to concentrate. The continuous measure of audit firm industry specialisation is described in Equation 8.[53]

INDSP_cont	=	the sum of the square root of the total assets of the clients that an audit firm services in a particular industry divided by the sum of the square root of the total assets of all clients of that audit firm.	(8)

Prior studies also use a dichotomous measure of audit firm industry specialisation (Krishnan 2001; Payne 2008). Krishnan (2001) argues that, in the absence of audit firm industry specialisation, an audit firm's portfolio share should be distributed evenly over the industries (i.e. 1 / number of industries). If an audit firm chooses to concentrate their operations in particular industries, this strategy will be reflected in a portfolio share of greater than 1 / number of industries. Payne (2008) applies Krishnan's approach, but argues that the benchmark should be adjusted according to

[52] In my robustness tests, I also used the alternative measures of audit firm industry specialisation described in Chapter 4. I present and analyse these results in Chapter 7.

[53] I also re-estimated my tests using specialisation metrics based on clients' total sales (as per Krishnan 2003; Balsam et al. 2003) and audit firms' total audit fees (as per Habib and Bhuiyan 2011; Bhattacharya 2011). In these untabulated regressions, I obtained qualitatively similar results to my main tests.

the sample to bring the portfolio-share measure more in line with the market-share approach; that is, to avoid recognising a substantially higher percentage of industry portfolio specialists than industry market specialists. Payne (2008) uses a dichotomous measure of industry specialisation that defines a specialist as existing when the continuous measure of portfolio share is greater than 3 / number of industries (i.e. a particular industry is over-represented in an auditor's portfolio by a factor of 3). The advantage of the dichotomous measure over the continuous measure is that it only recognises dominant auditors as being specialised. It is possible that industry expertise accrues only when industry portfolio share reaches certain threshold levels (Balsam et al. 2003). However, the empirical usefulness of the dichotomous measure of specialisation depends on the suitability of the arbitrary cut-off threshold chosen. My primary tests use Payne's benchmark to derive the dichotomous measure of portfolio specialisation, as per Equation (9):

INDSP_dum = 1 if *INDSP*_cont > (3 / number of two-digit industry codes used in the analysis in any given year),[54] 0 otherwise.[55] (9)

In my additional analysis, I test the sensitivity of my main results to variations in the definition of the cut-off threshold above.

[54] I also test alternative thresholds because the threshold at which the dichotomous variable for industry specialisation defined is arbitrary; I report and analyse the results in Chapter 7. BCK do not use a dichotomous measure of industry specialisation in their tabulated regressions.

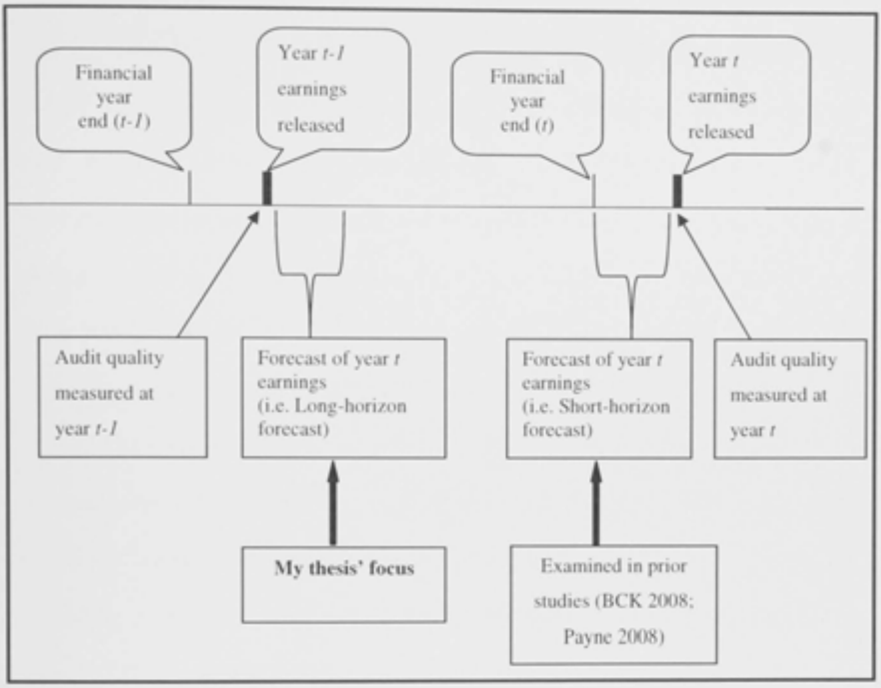
[55] There are between 52 and 59 two-digit SIC industries included in the analysis in each year within my sample period. Although Payne (2008) aggregates observations in three multi-year periods for estimating specialisation variables, I estimate the industry specialisation measures annually, to enhance the comparability with BCK's results. Untabulated regressions using Payne's pooled method generate similar results to those reported in my thesis.

5.3.3 Relative Timing of the Measurement of Forecast Accuracy and Audit Quality

The long-horizon and short-horizon regressions estimated in my thesis measure both audit firm industry specialisation and forecast accuracy at different times within the reporting year. Figure 5.1 illustrates the typical timing of long- and short-horizon forecasts and their relation to the measurement of audit firm industry specialisation. When testing H1a, which examines the accuracy of short-horizon forecasts, I measure audit firm industry specialisation in the year of the earnings that is the subject of the forecasts (year *t*), consistent with the established literature (Payne 2008; BCK 2008).[56] When testing H1b, which examines long-horizon forecast accuracy, I measure audit firm industry specialisation in the year prior to the earnings being forecast (year *t-1*). Since long-horizon accuracy is argued to reflect the quality of already published financial reports, audit firm industry specialisation (audit quality) is measured contemporaneously with those reports.

[56] BCK (2008) do not identify clearly the timing of their audit quality measurement. I inferred the timing of their measurement from their theoretical discussions of the reasons for their hypotheses.

Figure 5.1 Illustrations of the Forecast Errors and Audit Quality Being Examined



5.3.4 Proxy for Firm's Operating Risk

Hypothesis 2 proposes that audit firm industry specialisation has a greater effect on earnings quality when the firm's operating risk is higher, and that analyst forecast accuracy is more strongly associated with audit quality in such cases.

I use cash flow volatility as my main proxy for client operating risk, not only because this measure captures client's underlying riskiness (Minton et al. 2002; Allayannis et al. 2005), but also because cash flow volatility is largely exogenous to audit quality.[57] Cash flow volatility is measured as the natural logarithm of the five-

[57] I also use the innate accrual quality measure, developed by Francis et al. (2005a) to proxy client firm's operating risk in robustness tests. I discuss the details in Chapter 8.

year standard deviation of net cash flows from operating activities deflated by average total assets.[58] Net cash flow from operations has been shown to be more persistent than accruals and subject to less management manipulation (Sloan 1996; Xie 2001). Conversely, earnings volatility, which is affected by the behaviour of both cash flows and accruals, reflects economic volatility as well as the impact of managerial accrual decisions, which may be subject to discretionary biases such as income smoothing (Levitt 1998; Leuz et al. 2003) and 'bath-taking' (Givoly and Hayn 2000). Audit quality may be causally related to clients' earnings volatility, as high-quality auditors may be more effective in constraining accrual manipulations. Conversely, reported operating cash flow is relatively independent of audit quality, because auditors have no obvious responsibility to discipline the cash transactions in which the firm engages, and these transactions are relatively objective and verifiable. Therefore, consistent with prior literature (Dichev and Tang 2009; Jayaraman 2008), I use cash flow volatility to proxy firm's underlying riskiness.

5.3.5 Proxies for Analyst Quality

Hypotheses 3a and 3b predict that audit firm industry specialisation has a greater impact on the accuracy of lower-quality analysts. High-quality analysts are argued to have access to more private information and to possess superior ability to process complex information. As such, they are able to make more accurate forecasts than are other analysts. As analyst quality cannot be observed directly, I use various proxy measures drawn from prior literature, including analysts' general and firm-specific experience, the size of the analyst's employer (brokerage size) and an 'All-Star' measure. As discussed in Chapter 4, both the extant theory and empirical

[58] The natural logarithm of cash flow volatility is more normally distributed than the raw measure, as are the resulting regression residuals.

findings suggest that the direction of any relationship between analysts' portfolio complexity and analyst forecast accuracy is not clear. Thus, I include analysts' portfolio complexity as a control variable only, to reduce potential omitted variable bias. The measurement of this control variable is described in Section 5.3.6.

At any given time t, the quality of an individual analyst i, is estimated by the following measures:

General experience$_{i,t}$	=	the number of years through year t for which analyst i supplied at least one forecast for any firm;
Firm-specific experience$_{i,j,t}$	=	the number of years through year t for which analyst i supplied at least one forecast for firm j;
Brokerage size$_{i,k,t}$	=	the number of analysts employed by brokerage k, which employs analyst i in year t; and
All-Star$_{i,t}$	=	an indicator variable equal to '1' if analyst i was ranked by *II's* All-America Research Team in year t.

As H3a makes predictions regarding firm-year relationships, the firm-year variants of the analyst quality proxies described above are used to test that hypothesis. However, as H3b predicts differences in the performance of the individual analysts following a firm in a given year, analyst quality is measured at the analyst-firm-year level when testing this hypothesis. These firm-year and analyst-firm-year analyst quality measures are described in detail below.

To test H3a, which predicts that the average quality of analysts following a firm moderates the negative association between analysts' absolute forecast errors and audit quality, forecast error is regressed on audit quality, each of the analyst quality proxies and their interactions with audit quality (Equation 3). The mean value of the analyst quality proxies is computed at firm-year level within the forecast window.

For example, if a firm is followed by 10 analysts during the long-horizon forecast window in year *t*, individual analysts' general experience is first calculated and the average of these statistics is used as the measure of the test variable. Similar calculations apply to other analyst quality proxies.

I also use two composite measures based on the specific analyst quality measures, as illustrated in Table 5.1. The first composite measure (*CSCORE1*) is a function of all four analyst quality proxies, while the second measure (*CSCORE2*) includes only those proxies that capture analysts' personal attributes (general and firm-specific experience and 'All-Star' status).[59] To construct the composite score measures, I first identify all of the N analysts who submitted a forecast for any firm during a particular firm's beginning-of-year forecast window. I then rank analysts according to each individual quality proxy. If there are N analysts issuing forecasts, the lowest analyst ranking for a quality proxy is ranked 1, and the best analyst is ranked N. In cases in which there is a tie, the average of the adjacent ranks is assigned to these analysts. A stylised example of these calculations is provided in Table 5.1. There are five analysts (A, B, C, D and E) active during Firm X's beginning-of-year forecast window, among which Analysts D and E have equivalent 'firm experience' (Panel A). In this case, the average of the adjacent ranks (3.5) is assigned to these analysts. A related procedure is applied to the ranking for 'All-Star' status: all 'All-Star' analysts and all non-star analysts are assigned average ranks. For example, Analysts C and D are 'All-Star', and thus their average ranks are 4.5 (i.e. [5+4]/2 = 4.5). The average rank for the non-star analysts (Analysts A, B and E) is 2. I then sum the

[59] I exclude brokerage size in the second composite measure for two reasons. First, it is a proxy capturing analyst employers' attributes, rather than the quality of an individual analyst. Second, a recent study finds no relationship between brokerage size and forecast accuracy in the post-Reg FD period (Keskek et al. 2013). Therefore, inclusion of brokerage size in the calculation of the composite score may contaminate the results (bias against finding a result).

individual quality ranks for each analyst and divide this value by the number of analysts covering the firm. For example, Analyst A has a total rank across all four proxies of 15 (5+5+3+2), and a total rank excluding that pertaining to brokerage size of 12 (5+5+2). In Panels B and C, Firm X is followed by Analysts A, B and C, whose rank sums are 15, 12 and 11.5 (12, 8 and 6.5 if brokerage size is excluded); the average total ranks for analysts following Firm X are 12.83 ([15+12+11.5]/3) for *CSCORE1* and 8.83 for *CSCORE2*. The average per analyst adjustment allows for valid comparison across firms that vary in the level of analyst coverage.

Table 5.1: An Illustration of the Procedures to Generate the Composite Score

Panel A: Analysts' Rank within the Long-Horizon Forecast Window of Year 2000

	Firm Following	General Experience	**Rank**	Firm Experience	**Rank**	Brokerage Size	**Rank**	All-Star	**Rank**	**Sum 4 Ranks**	**Sum 3 Ranks**
Analyst A	**X**	8	**5**	6	**5**	30	**3**	0	**2**	**15**	**12**
Analyst B	**X**	7	**4**	3	**2**	40	**4**	0	**2**	**12**	**8**
Analyst C	**X**	1	**1**	1	**1**	50	**5**	1	**4.5**	**11.5**	**6.5**
Analyst D	**Y**	4	**2**	4	**3.5**	10	**1**	1	**4.5**	**11**	**10**
Analyst E	**Y**	6	**3**	4	**3.5**	20	**2**	0	**2**	**10.5**	**8.5**

Panel B: Composite Score 1 (4 proxy) for a Given Firm in Year 2000

Sum 4 Ranks	Analyst A	Analyst B	Analyst C	Analyst D	Analyst E	*CSCORE1*
Firm X	15	12	11.5			**12.83**
Firm Y				11	10.5	**10.75**

Panel C: Composite Score 2 (3 proxy) for a Given Firm in Year 2000

Sum 3 Ranks	Analyst A	Analyst B	Analyst C	Analyst D	Analyst E	*CSCORE2*
Firm X	12	8	6.5			**8.83**
Firm Y				10	8.5	**9.25**

Tests of H3b regress the difference in the forecast accuracy between the 'worst' quality and 'best' quality analyst (estimated separately for each of the individual analyst quality proxies) on audit firm industry specialisation and the control variables, as per Equation (5). As each of my continuous proxies for analyst quality is increasing in their expected effect on analyst performance, the 'worst' ('best') quality analyst is identified as the analyst who has the minimum (maximum) value of the particular proxy in question. For example, if a firm is followed by several analysts in year *t*, with their general experience ranging between two and eight years, the analyst with two years' experience is identified as the 'worst' and the analyst with eight years' experience is identified as the 'best' quality analyst. If the 'worst' quality analyst has an absolute forecast error of \$0.45, and the 'best' quality analyst has an absolute forecast error of \$0.20, the difference in the absolute forecast errors between the 'worst' and 'best' quality analysts is \$0.25 (\$0.45 – \$0.20). This difference is then regressed on audit firm industry specialisation and the control variables to test H3b. Where ties occur with respect to the identification of the 'worst' ('best') analyst, the mean value of the absolute forecast errors of these 'worst' ('best') analysts is used. This calculation is repeated separately for each continuous analyst quality proxy (i.e. firm experience and brokerage size). For tests involving the dichotomous 'All-Star' status proxy, the mean value of the absolute forecast errors of non-star analysts following a firm is computed from which the mean value of the absolute forecast errors of 'All-Star' analysts is subtracted. Tests using the 'All-Star' proxy are only possible for cases where there are at least one 'All-Star' and one non-star analyst following a firm.

Similar to the tests of H3a, I construct two composite score measures for testing H3b. The first composite score (*CSCORE3*) is calculated by summing the analysts' rankings for all four analyst quality proxies where the ranking is conducted within the cohort of analysts following a firm in a given year, rather than summing across all analysts active within the long-horizon forecast window. Accordingly, the second measure (*CSCORE4*) excludes the rankings for brokerage size. This difference in the calculations of composite score measures for testing H3a and H3b reflects the fact that H3b examines the relative performance of the 'worst' and 'best' quality analysts following a firm in a given year. Thus, within-firm rankings are required to identify the 'worst' and 'best' analysts following a firm. An example of this calculation is shown in Table 5.2. If the *CSCORE3* (based on all four specific analyst quality proxies) is considered, Analyst A has the highest within-firm-year composite rank (8.5) and is thus identified as the 'best' analyst, while Analyst B has the lowest within-firm-year rank (7.5) and is thus identified as the 'worst' analyst.[60] When the 'worst' and 'best' analysts are identified based only on analysts' personal attributes (*CSCORE4*), Analyst A is still the 'best' analyst of Firm X, but Analyst C is now the 'worst' analyst.[61]

---

[60] If I use the ranking procedure (i.e. ranking within the long-horizon window) employed for the tests of H3a, Analyst A is still the 'best' analyst, but Analyst C is now the 'worst' analyst (see Table 5.1).

[61] Since the 'All-Star' proxy is an indicator variable, an 'All-Star' analyst will receive a much larger ranking value when a firm is followed by a relatively large number of analysts where most of them are non-stars. Therefore, in the sensitivity tests, I used different ranking procedures for 'All-Star' status. For example, I assign 'All-Star' analysts a ranking value of 3, 5 or 10, while assigning non-star analysts a ranking value of 1, or exclude the 'All-Star' proxy in the calculation of the composite measures. I discuss the results in Chapter 8.

Table 5.2: Analysts' Rank within a Firm of Year 2000 (*CSCORE3* and *CSCORE4*)

	Firm follows	General Experience	**Rank**	Firm Experience	**Rank**	Brokerage Size	**Rank**	All-Star	**Rank**	**Total Rank (*CSCORE 3*)**	**Total Rank (*CSCORE 4*)**
Analyst A	**X**	8	**3**	6	**3**	30	**1**	0	**1.5**	**8.5**	7.5
Analyst B	**X**	7	**2**	3	**2**	40	**2**	0	**1.5**	**7.5**	5.5
Analyst C	**X**	1	**1**	1	**1**	50	**3**	1	**3**	**8**	5
Analyst D	**Y**	4	**1**	4	**1.5**	10	**1**	1	**2**	**5.5**	4.5
Analyst E	**Y**	6	**2**	4	**1.5**	20	**2**	0	**1**	**6.5**	4.5

5.3.6 Control Variables

Four of the hypotheses (H1a, H1b, H2 and H3a) proposed in my thesis are tested using the same dependent variable, analysts' absolute forecast errors. I control for several variables identified in the prior literature as being associated with analysts' forecast errors, and which are also plausibly correlated with my test variables. While most of the control variables are common, I include additional control variables in the tests of H3a, which examines the impact of analyst quality on the association between audit firm industry specialisation and forecast accuracy. These additional variables are likely to be correlated with the analyst quality proxies, and their omission may bias the estimated regression coefficients for my test variables. The tests of H3b use the difference in forecast accuracy between the 'worst' and 'best' quality analysts as the dependent variable. In addition to the control variables common to other models, I include variables that are logically correlated with the difference in forecast accuracy across analysts of different quality (e.g. measures of the magnitude of the difference in analyst quality). Below, I discuss the common control variables (Section 5.3.6.1), followed by a discussion of the additional control variables used to test H3a (Section 5.3.6.2) and H3b (Section 5.3.6.3).

5.3.6.1 Common control variables for all tests and predictions

This section describes the control variables common to all tests. The predicted signs for the control variables described are applicable to the tests of H1a, H1b, H2 and H3a (where analysts' absolute forecast error is the dependent variable). While I also include theses controls when testing H3b (where the difference in forecast errors across analysts is the dependent variable), their directional impact may differ from that predicted for the earlier hypotheses. I discuss the reasons for this at the end of the current section.

I include analysts' forecast dispersion (*DISP*) to control for the impact of analysts' broader information environment on their forecast accuracy. Analysts may have access to different levels of private information about firms' future earnings, which are reflected in the forecasts they issue. Forecast dispersion is found to be positively related to forecast errors, as firms' information uncertainty makes the forecasting task more complex, leading to decreased forecast accuracy (Lang and Lundholm 1996; Zhang 2006). I measure *DISP* as the standard deviation of analysts' earnings forecasts, deflated by beginning-of-month stock price. A positive coefficient is predicted between *DISP* and forecast errors. Forecast dispersion is not included in my tests of H2, which examines the impact of firms' underlying riskiness on the relation between audit firm industry specialisation and forecast accuracy. Unlike cash flow volatility, which largely captures firms' underlying riskiness, forecast dispersion reflects uncertainty related to both the volatility of a firm's underlying fundamentals and its information quality (which may be endogenous to the auditor's identity). In my sensitivity analysis, I re-estimate my tests of H2 within sub-samples defined by the level of forecast dispersion.

The forecast horizon (*HORIZON*) is also controlled, measured as the natural logarithm of the number of days elapsing between the date on which a forecast is issued and the earnings announcement date. As the set of information available to analysts increases as the earnings announcement date approaches, forecast errors decrease in parallel. Therefore, the predicted direction for the relation between absolute forecast errors and the forecast horizon is positive.

Consistent with prior literature, I control for the size of the covered firm (*SIZE*). To the extent that size proxies the firm's information environment, a negative relationship between firm size and absolute forecast errors is expected. This is because larger firms' access to more available information facilitates greater accuracy in their earnings forecasts (Atiase 1985; Collins et al. 1987). Conversely, other studies argue that size reflects stronger managerial incentives or greater financial reporting complexity, which may increase forecasting task difficulty, and show that size is positively correlated with analysts' absolute forecast errors (Hope 2003; Dhaliwal et al. 2012). For these reasons, I do not predict a direction for firm size. Firm size is measured as the natural logarithm of the firm's total assets or market value of the equity at the beginning of the year (Dhaliwal et al. 2012; Payne 2008; BCK 2008).

Analyst following (*NUMEST*), measured as either raw or the natural logarithm of the number of analysts covering a firm, is predicted and found to be negatively correlated with absolute forecast errors (Lys and Sohn 1990; Payne 2008). Greater analyst following imposes pressure on firms for supplementary disclosures, increases the competition among analysts and promotes analysts' incentives for accuracy.

Similarly, it is contended that firms subject to greater uncertainty, such as those with larger absolute accruals and those suffering greater financial distress, have larger analysts' absolute forecast errors (Lang and Lundholm 1996; Payne 2008; BCK 2008). Firms' absolute accruals (*ABSACCR*) are measured as the difference between net income and cash flow from operations,[62] while financial distress is proxied by Zmijewski's (1984) *ZSCORE*.[63] Positive coefficient estimates are predicted for these two variables.

Prior research also finds that analysts' absolute forecast errors are greater for loss-reporting firms than for profit-reporting firms (Das et al. 1998; Gu and Wu 2003). This is because loss-reporting firms have relatively uncertain earnings, and analysts may not be able to assess fully the performance or transitoriness of the losses, increasing the likelihood of absolute forecast errors. An indicator variable (*LOSS*) is included to control for this effect and the coefficient sign is predicted to be positive.

Prior earnings are typically a strong predictor of future earnings. When firms experience a large earnings change in the current year, or have greater standard deviation of return on equity in prior years, their past earnings are a relatively weak indicator of future earnings, and analysts' forecast errors are typically larger (Kross et al. 1990; Lang and Lundhom 1996; Dichev and Tang 2009). Therefore, the absolute value of firms' annual earnings change (*ABSECHG*) and standard deviation of return on equity over the previous five years (*STDROE*) are controlled. The

---

[62] Firms' absolute accruals (*ABSACCR*) are very likely to be endogenous to auditor identity. Therefore, I only include this variable in tests of H1a and H1b to maintain consistency with Payne (2008).

[63] Zmijewski's (1984) financial distress ratio measures firm performance, leverage and liquidity. The ratio is developed based on 40 bankrupt and 8000 non-bankrupt firms: ZSCORE = -4.336 - 4.513*ROA + 5.679*leverage – 0.004*liquidity.

coefficients for these two variables are expected to be positive. I also include a control for earnings level (*EL*) because prior research demonstrates a positive relationship between this measure and absolute forecast errors (Eames and Glover 2003; BCK 2008). Other studies show that forecast pessimism (optimism) increases when the level of a firm's earnings increases above (declines below) the average earnings of all firms (Hwang et al. 1996; Eames and Glover 2003). Thus, a positive relationship is predicted for *EL*. Following Payne (2008), a dichotomous measure of earnings persistence identifying firms in the middle three quintiles of the signed change in earnings (*PERSIST*) is included to control for threshold effects in persistence. Firms for which *PERSIST* equals 1 are expected to have smaller forecast errors, and the coefficient sign is predicted to be negative.

I also include the control variables identified above in the tests of H3b, in which the dependent variable is the difference in forecast errors of the 'worst' and 'best' quality analysts within a given firm. As no prior research employs this dependent variable, I develop and justify predicted signs for the control variable coefficients logically, relying on similar assumption to those underpinning H3b. If the control variables capture factors related to an increase in the forecasting task complexity (increase earnings uncertainty or firms' overall information uncertainty), the positive impact of these variables on absolute forecast errors is likely to be greater for the 'worst' quality analyst than for the 'best' quality analyst. Therefore, *DISP*, *ZSCORE*, *LOSS*, *ABSACCR*, *ABSECHG*, *STDROE* and *EL* are likely to be positively correlated with the difference in absolute forecast errors between the 'best' and 'worst' analysts following a firm. Likewise, if the control variables proxy factors that decrease forecasting task complexity, they should have a greater impact in reducing the absolute forecast errors of the 'worst' analyst compared to those of the 'best' analyst. Thus, negative coefficients are predicted for these variables (*NUMEST* and *PERSIST*). I also include additional variables in testing H3b to capture the analysts' relative performance drivers. These variables are detailed in Section 5.3.6.3.

5.3.6.2 Additional control variables for tests of H3a

H3a predicts that audit firm industry specialisation has a greater impact on the accuracy of forecasts made by lower average quality analysts following a firm in a given year. In testing this hypothesis, I include proxies for the mean quality of analysts following a client firm in a particular year, and the interactions between these proxies and audit quality. There are two reasons that I include additional control variables to capture analysts' portfolio complexity. First, Clement (1999) argues that portfolio complexity may be correlated with analyst quality variables such as general experience because more experienced analysts typically follow portfolios that are more complex. Further, since analysts employed by larger brokers typically follow less complex portfolios, portfolio complexity may be correlated with brokerage size. Second, there is evidence, albeit of an inconsistent direction, that analysts' portfolio complexity is correlated with analyst forecast accuracy (Clement 1999; Kim et al. 2011; Drake and Myers 2011). Therefore, I control for portfolio complexity when testing H3a to avoid potential omitted variable bias. Following prior studies (Clement 1999; Jacob et al. 1999; Drake and Myers 2011), I use two measures for analysts' portfolio complexity: the number of firms following (*FFOLLOW*) and the number of industries following (*IFOLLOW*). I measure *FFOLLOW* as the number of firms an analyst follows over the forecast window in year t, and *IFOLLOW* as the number of industries an analyst follows over the forecast window in year t, where the number of industries is calculated as the number of two-digit SICs.[64] Recall that portfolio complexity may be negatively correlated with forecast accuracy due to the reduced time and attention analysts devote to each firm and lesser likelihood of an analyst obtaining private information from management. Conversely, portfolio complexity may be positively correlated with forecast accuracy if analysts benefit more significantly from knowledge spillovers across covered firms and industries. Therefore, there is no predicted sign for these two portfolio complexity measures.

5.3.6.3 Additional control variables for tests of H3b

H3b predicts that audit firm industry specialisation reduces the difference in forecast accuracy between the 'worst' and 'best' analysts following a firm in a given year. In testing H3b, the dependent variable is the difference in the forecast accuracy between the 'worst' and 'best' quality analysts (estimated separately for each of the individual analyst quality proxies at the analyst-firm-year level). Thus, when testing H3b, I include five additional control variables (*DIFANQ*, *ANQ_B*, *ABSFE_B*, *HORIZON_B* and *DIFHORIZON*) in an attempt to capture the effect of the extent to which the 'best' and 'worst' analysts differ in quality, and the extent to which the timing of their forecasts differs. Since there are no prior studies that have used my dependent variable, I select my control variables based on their intuitive correlation with the dependent variables and other test variables.

---

[64] I also estimate regressions for which portfolio complexity is measured based on analysts' coverage of firms over the entire year t, rather than over the forecast window. For example, *FFOLLOW* is the average of the number of firms covered by each analyst who issues a forecast for firm j during the entire financial year. The untabulated results are similar to those reported in Chapter 8.

I include the difference in the actual value of the quality proxy (*DIFANQ*) between the 'best' and 'worst' analysts (e.g. if the 'worst' analyst has three years' general experience while the 'best' analyst has 10 years' experience, *DIFANQ* equals 7). As previously discussed, the greater accuracy of the forecast of the 'best' analyst relative to the 'worst' analyst is attributable to the 'best' analyst's access to a greater amount of private information and superior ability to analyse this information. Inclusion of *DIFANQ* captures the degree of disparity in analyst quality and its impact on the relative performance of analysts. A negative coefficient is thus expected. Since the impact of *DIFANQ* may be contingent on the quality level of the 'worst' (or 'best') analyst identified, I also control for the level of the quality proxy for the 'best' analyst (*ANQ_B*). For example, a difference in general experience of three years may have negligible empirical meaning if the 'best' analyst has 15 years' experience, but be quite important if the 'best' analyst has five years' experience. *ANQ_B* may also be systematically correlated with the dependent variable. When the 'best' analyst's ability to forecast increases, the forecast is likely to be more accurate, reducing the difference in forecast accuracy between this analyst and the 'worst' analyst. A positive coefficient is thus predicted.

I also control for the actual level of absolute forecast errors of the 'best' quality analyst (*ABSFE_B*), to capture the underlying level of forecast complexity. I predict a negative coefficient for *ABSFE_B* because *ABSFE_B* captures the underlying forecasting difficulty pertaining to all analysts (e.g. the difficulty to access a greater amount of private information). When this difficulty increases, analyst forecast accuracy should decrease and the impact is likely to be greater for the 'best' quality analyst than for the 'worst' quality analyst. Therefore, the difference between the forecast accuracy between the 'worst' and 'best' analysts decreases, and I expect *ABSFE_B* to have a negative coefficient.[65]

In the tests of H3b, I replace the average forecast horizon of analysts following a given firm with the forecast horizon of the 'best' analyst (*HORIZON_B*) and add a control for the difference in the forecast horizon between the 'worst' and 'best' analysts (*DIFHORIZON*).[66] As argued earlier, forecast errors decrease when the timing of forecast approaches the earnings announcement date. While the average forecast horizon (*HORIZON*) is included in models examining the consensus forecast accuracy of a firm, the forecast horizon of an individual analyst (*HORIZON_B*) should logically be used in models assessing the individual analyst forecast accuracy of a given firm. Similar to the *HORIZON* variable, a positive coefficient for *HORIZON_B* is predicted. I also include the difference in the forecast horizon (*DIFHORIZON*) between the 'worst' and 'best' analysts to control for the impact of timing difference between these analysts on their relative performance. If *DIFHORIZON* increases (the 'worst' analyst's forecast is issued earlier than the 'best' analyst's forecast), the absolute forecast error of the 'worst' analyst is likely to be greater than that of the 'best' analyst because the older forecast is issued using a less complete information set. Therefore, I expect *DIFHORIZON* to have a positive coefficient.

---

[65] In my sensitivity tests, I re-estimate my models by dropping *ABSFE_B*, replacing it with the average absolute forecast errors of the 'worst' and 'best' analysts (*ABSFE_AVG*), or with the absolute forecast error of the 'worst' analyst (*ABSFE_W*). My main results are not affected.

[66] When ties occur with respect to the identification of the 'worst' ('best') analyst, the mean value of actual quality proxy of the 'worst' ('best') analysts as well as their forecast horizons are used.

5.4 Estimation Methods

Section 5.2 introduced the general form of the regression models used to test my hypotheses and identified the reasons that the standard OLS regression applied to these models may generate statistically inconsistent estimators. This section introduces and explains the three main estimation methods used in my thesis: OLS regression with clustered standard errors, two-stage regressions (2SLS or Heckman treatment regression) and a two-stage matched-sample approach (PSM regressions). Further, the extent to which each method is likely to produce efficient and consistent estimators in the presence of the validity threats identified in Section 5.2 is explained.

5.4.1 Ordinary Least Squares with Clustering

OLS standard errors may be biased for samples comprising panel data. The adjustment of standard errors for clustering is a response to the potential bias present when residuals are correlated across observations. As my samples consist of multiple observations for the same firms, adjusting standard errors for the observed within-firm correlation in errors produces theoretically superior test statistics for regression coefficients. Therefore, I estimate models using OLS with standard errors adjusted for within-firm clustering as per the method of Petersen (2009). The Petersen one-way clustering method is appropriate when residuals are correlated within firms but not years. For my tests of H1a, H1b, H2 and H3b, I control for potential time effects by including year dummies, and thus do not cluster standard errors by years. However, where the regressors of interest (test variables) include significant time-related effects, using year dummies may artificially inflate standard errors, and a two-way clustering by firm and year reduces bias and is most appropriate (Thompson 2011). My tests of H3a examine the relation between audit firm industry specialisation, analyst quality and forecast accuracy where several measures of analyst quality are collinear with the year dummies. This collinearity arises largely by construction: for example, the average analysts' general and firm-specific experience increase with the age of my sample, as experience can only be measured within the period for which both analysts and covered firms are included in the *I/B/E/S* database.[67] Thus, clustering by firm and year rather than including year dummies is desirable for testing H3a.

5.4.2 Two-stage Regression Approaches (2SLS or Heckman)

Section 5.2 identified the possibility that the selection of an industry specialist auditor is determined endogenously, which may result in biased OLS estimates of the association between forecast accuracy and audit firm industry specialisation. The use of two-stage regressions may mitigate the limitations of OLS under these conditions. Below, I describe the two-stage approaches appropriate to continuous and dichotomous test variables: 2SLS and Heckman treatment effects regressions. While the use of Heckman regressions is restricted to my robustness tests, it is efficient to describe their nature here.

The generic two-stage approach regresses the endogenous variable (industry specialisation) against the explanatory variables in a first-stage model. Statistics generated by the first-stage regression are then used to control for the impact of endogenous selection in the second-stage regression (Wooldridge 2008). The specific approach adopted varies according to whether the endogenous variable is measured continuously or dichotomously. The 2SLS method is applied when

---

[67] This induced collinearity is not of concern in my tests of H3b because the experience measures are simply used to rank analysts.

industry specialisation is measured by a continuous variable. Under 2SLS, the first-stage regression is estimated by OLS with the predicted value for industry specialisation obtained from this regression subsequently replacing the actual value of industry specialisation in the second-stage model (Wooldridge 2008, 521). For the dichotomous measure of industry specialisation, the Heckman treatment effects method may be used. Here, the first-stage model estimates the likelihood of appointing a specialist using probit regression. For each observation, the inverse Mills ratio (IMR), which is a decreasing function of the probability of selection, is estimated, and this is included as an additional regressor in the second-stage model to control for the endogenous selection of auditor. Subject to satisfaction of specification tests (described below), the coefficient for the endogenous variable (industry specialisation) can be estimated without bias (Lennox et al. 2012). When using 2SLS or Heckman regressions, it is necessary to identify statistically valid instrumental variables, which are included in the first-stage, but not in the second-stage regression (Wooldridge 2008, 521; Lennox et al. 2012).[68] The general form of the first-stage regression is illustrated in Equation (10).

$$INDSP = \beta_0 + INSTRUMENTAL\ VARIABLES + CONTROLS + \varepsilon \quad (10)$$

There are two important concerns regarding the use of two-stage regressions: consistency and efficiency. The two-stage approaches employed here will only produce statistically consistent estimators if the instrumental variables used satisfy relevance and exogeneity conditions. Where consistent two-stage estimators are

[68] While not recommended, Heckman regressions can be estimated without the use of external instruments because the IMR acts as an exclusion restriction. However, the use of the Heckman method without instruments requires the model to be identified solely on the distributional assumptions about the residuals (refer to Sartori 2003 for more details). Further, the Heckman model without instruments is more likely to be subject to multicollinearity problems, which may inflate the coefficient standard errors or bias the estimates if the model is also misspecified (Lennox et al. 2012). Thus, I apply the instruments to both the 2SLS and the Heckman regressions.

generated, these will typically be less efficient than would OLS estimators; thus, it is normally desirable to test the consistency of OLS estimators to determine the appropriate method for a particular case. These issues are now discussed in detail.

Valid instrumental variables must be both relevant to the explanation of the endogenous variable and exogenous to the errors in the second (structural) equation (these conditions are typically described as the 'relevance' and 'exogeneity' conditions). If the instrumental variable(s) included in the first-stage regression is (are) not, in fact, exogenous *and* relevant, the two-stage approach may produce biased estimators (Larcker and Rusticus 2010).

Instrument relevance is a function of the correlation between the endogenous regressor and the candidate instrument, and can be assessed by tests of under-identification or weak identification. The first-stage equation is regarded as under-identified if the number of instruments is less than the number of endogenous regressors, or in cases in which the former is greater than the latter but there is no correlation between them. In cases in which the number of instruments exceeds the number of endogenous regressors and there is only weak correlation between the instruments and endogenous regressors, the first-stage equation is described as weakly identified (Wooldridge 2008, 514–517). I begin by proposing two candidate instruments for a single endogenous regressor (audit firm industry specialisation). I then test the candidate instruments for under- or weak identification. The Lagrange multiplier (LM) version of the Kleibergen-Paap rk test assesses under-identification (Baum et al. 2007). This is essentially a test of whether the correlation between the instrument(s) and the endogenous regressor is statistically different from zero

(Kleibergen and Paap 2006). Weak instruments are identified using Wald-type F-statistics (Baum et al. 2007).[69] These tests simply examine the coefficients of the instruments in the first stage, and have a null hypothesis that the instrument(s) is (are) jointly zero. Thus, the null is rejected if the coefficient(s) of the instrument(s) is (are) significantly different from zero, suggesting that the equation is identified (Stock et al. 2002; Baum et al. 2007). Stock et al. (2002) propose benchmarks for determining weak instruments using F-statistics, which I consult in developing my models. Further, the inclusion of the partial effect of the instruments (i.e. partial R^2) may help to determine the presence of weak instruments (Larcker and Rusticus 2010).

An instrument is exogenous if it is strictly uncorrelated with the error term in the second-stage equation. In models where the number of instruments is equal to the number of endogenous variables, the exogeneity of instruments can only be evaluated by the examination of the theoretical justification for exogeneity (Wooldridge 2008, 529–530). In cases where there are more candidate instruments than the number of endogenous variables (i.e. the model is over-identified), specification tests, such as the Sargan-Hansen over-identification test, are available to assess formally the exogeneity of the instruments. Under the Sargan-Hansen test, the R^2 statistics are obtained from a model that regresses the second-stage residuals on all exogenous variables. The R^2 should be statistically indistinguishable from zero if the instruments are exogenous (Larcker and Rusticus 2010). The null hypothesis

[69] The LM test, which tests the improvement in model fit (i.e. R^2) by including additional variables (i.e. instrument) in the model, can be used to examine under- or weak-identification (Kleibergen 2007). However, Zivot et al. (1998) show that the LM test is superior to the Wald-test in the detection of weak identification only in a single instrument case, but not in the case of multiple instruments. I identify two instruments and thus use the Wald-type F-statistics to examine weak-identification.

under the Sargan test is that all proposed instruments are exogenous. As such, failure to reject the null corroborates the validity of the instruments.[70]

Conditional on the existence of valid instruments, the efficiency of the instrumental variables estimation procedure can be assessed via the Durbin-Wu-Hausman test. Hausman-type tests assess the statistical difference between a vector of consistent estimators (two-stage estimators) and a vector of efficient estimators (OLS) (Wooldridge 2002, 118–120). If the efficient estimators are not statistically different from the consistent estimators, the null hypothesis of the Hausman test cannot be rejected and the efficient estimator (OLS) is assumed to produce consistent estimates, and thus can be validly used in empirical tests. However, rejection of the null hypothesis under the Hausman test suggests that IV estimators are required for valid hypothesis testing because the effects of the endogenous regressors on the estimates are meaningful, and thus OLS estimators may be inconsistent (Larcker and Rusticus 2010).

5.4.3 Two-stage Matched-Sample Approach (Propensity Score Matching)

The previous section discussed the effectiveness of the two-stage approach in mitigating endogenous auditor choice. An alternative approach is the PSM method (Rosenbaum and Rubin 1983), which, like the Heckman treatment approach, potentially provides valid estimates of average treatment effects when the subject of interest (audit firm industry specialisation) is a dichotomous measure.

---

[70] I also estimate Sargan tests for models containing dichotomous measures of specialisation, because no equivalent test exists for Heckman-type models.

The PSM approach first computes the conditional probability (the propensity score) of a subject to a treatment based on a vector of variables, and then matches the treatment firms to control firms based on their propensity score, so that the differences in the relevant characteristics of the two groups are reduced. To test whether the matching procedure applied has generated a balanced sample, a univariate analysis of the mean differences in the matching variables across the treatment and control groups is conducted. If all of these mean differences are insignificantly different from zero, the sample is considered balanced with respect to these variables. A multivariate regression analysis can then be conducted to examine any causal relationships (Guo and Fraser 2010, 134).

In my thesis, a logistic regression is used to estimate the conditional probability of a client hiring an industry specialist auditor, with all control variables identified in Equation (10) included as predictors (similar to Minutti-Meza 2013).[71] I then attempt to match each treatment firm (clients audited by industry specialist auditors) to a control firm (clients audited by non-specialists) with a similar propensity score, subject to criteria determined by the properties of the sample and the balance test described above. The impact of audit firm industry specialisation on forecast accuracy is then examined by estimating an OLS regression on the propensity score matched sample.

Potential matching criteria include (but are not restricted to) nearest neighbour matching and nearest neighbour matching subject to a caliper (tolerance distance).

---

[71] For tests of H3b, I omit the control for the forecast accuracy of the 'best' analysts (*ABSFE_B*) from the first stage, as this is likely to induce a post-treatment bias. My results are not sensitive to this modelling choice.

The nearest neighbour method matches each treatment firm (clients audited by an industry specialist) to a control firm (clients audited by a non-specialist) with the smallest absolute difference in propensity score (Guo and Fraser 2010, 146). A limitation of this method is that it may potentially match clients whose propensity scores are very different. To mitigate this problem, a maximum tolerance distance, known as a caliper, can be applied to the nearest neighbour matched pairs, which filters pairs for which the absolute difference in propensity scores exceeds this tolerance distance. Rosenbaum and Rubin (1983) suggest the use of a caliper size less than or equal to a quarter of a standard deviation of the estimated propensity score of the sample. This matching approach has been widely used by researchers before conducting multivariate analyses (Guo and Fraser 2010, 149).[72] Saliently, Lawrence et al. (2011) apply the nearest neighbour matching within a caliper to examine the relation between audit firm size and short-horizon forecast accuracy, while Minutti-Meza (2013) applies the same matching approach in the examination of audit firm industry specialisation and other measures of audit outcomes (discretionary accruals, the auditor's propensity to issue a going-concern opinion and the client's propensity to meet or beat analysts' forecasts). Therefore, I also use this matching approach and follow the advice of Rosenbaum and Rubin (1983) to determine an appropriate caliper size.[73]

The PSM method has two potential advantages over the Heckman approach. First, it parses out the impact of client characteristics on auditor choice without assuming a

[72] Under each of the matching approaches described above, researchers must decide whether control firms can be used just once (matching without replacement) or several times in the sample (matching with replacement). To maintain consistency with the extant auditing literature (Lawrence et al. 2011; Minutti-Meza 2013), I use the matching 'without replacement' criterion.

[73] Several tolerance calipers are applied when using this criterion. The broadest caliper level that generates a matched sample and as such satisfies tests for balance across the treatment and control groups is used in my final analysis.

specific functional form, thus providing a more direct estimate of the treatment effects. In addition, when the underlying functional form is nonlinear, the matching approach reduces the effects of nonlinearities when estimating the treatment effects (Lawrence et al. 2011). Therefore, when the audit firm industry specialisation is a dichotomous measure, I use the PSM method in my main results tables. However, I report the results of the Heckman regression models in my additional tests.

Alongside the advantages, there are recognised limitations to the PSM approach. First, Cram et al. (2009) argue that results obtained from the reduced matched sample may not be generalisable to the full population; rather, they may only apply to the range of common support. To explore the degree of overlap in the propensity scores between the treatment and control firms (range of common support), I plot the histograms for the propensity scores of the two groups of firms after each matching, to observe visually whether the control firms have propensity scores spanning the full range of the propensity scores of the treatment firms. I then examine the pseudo R^2, the receiver operating characteristic curve (ROC) and the area under the ROC curve to evaluate the predictive power of the matching models. If the R^2 and the area under the ROC curve approach 1, the matching models are said to be of greater predictive power (Minutti-Meza 2013, 12). Second, matching is conducted on post-treatment attributes, which may introduce a bias in the results if the matching variables are affected by the auditor choice (Lawrence et al. 2011). In my main test, I use the matching according to propensity scores generated by the first-stage models described above, to be consistent with recent auditing studies (Lawrence et al. 2011; Minutti-Meza 2013) and to preserve the post-matching sample size. However, to address the possibility of bias induced by post-treatment matching, I re-estimate the models using the lagged values of the variables that are likely to be affected by auditor choice. In addition, because auditor identity is likely to be sticky, I replace the covariates potentially subject to a post-treatment bias with alternative variables that are theoretically independent of auditor identity.

5.5 Regression Models

The general form of the models used to test my hypotheses was described in Section 5.2. In this section, I describe and discuss in detail each of the empirical models used to test the hypotheses.

5.5.1 Models to Test Hypotheses 1a and 1b

To test H1a, I first estimate a series of regressions of short-horizon forecast accuracy on audit firm industry specialisation, using specifications based on those employed by BCK (2008) and Payne (2008), which generated contradictory results regarding the impact of industry specialisation on forecast accuracy. To reconcile these prior findings, I test the sensitivity of the base models to the use of alternative deflators for the dependent variable, additional control variables and endogeneity correction. I repeat this analysis using long-horizon forecast data to test H1b. All models employ a dependent variable, which is an inverse function of analyst forecast accuracy.[74]

5.5.1.1 Tests of H1a—Models based on existing literature

I first detail the similarities and differences in the models that have been applied in the existing literature, followed by a discussion of their limitations.

[74] The analysts' absolute forecast errors are measured as per Equations (6a) and (6b). The use of forecast errors corresponds to Payne's (2008) measurement, but contrasts with that of BCK (2008) who use the negative of their absolute forecast errors when measuring accuracy. Some variables have been renamed to enhance comparability across models.

BCK's Model

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 HORIZON + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ZSCORE + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 STDROE + \beta_9 EL + YEAR + \varepsilon$$

Payne's Model

$$ABSFE\_un = \beta_0 + \beta_1 INDSP + \beta_2 DISP\_eps + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ABSACCR + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 PERSIST + YEAR + INDUSTRY + \varepsilon$$

Where

ABSFE	=	analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price (an inverse function of forecast accuracy);
ABSFE_un	=	analysts' absolute forecast errors (undeflated), measured as the absolute value of the difference between actual *I/B/E/S* earnings per share and forecast earnings per share (an inverse function of forecast accuracy);
INDSP	=	the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm;
DISP_eps	=	forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the absolute value of the mean EPS forecast during the period;
HORIZON	=	the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date;
SIZE	=	the natural log of the market value of equity (natural log of total assets in Payne's model);
NUMEST	=	the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting (BCK model takes log of this number);
ZSCORE	=	Zmijewski's financial distress score;
LOSS	=	1 if a firm reports negative earnings, 0 otherwise;
ABSECHG	=	the absolute value of the change in annual earnings, deflated by opening stock price (natural log of total assets in Payne's model);
STDROE	=	the standard deviation of return on equity over the previous five years;
EL	=	earnings per share, winsorized at 5 (-5);
ABSACCR	=	absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; and
PERSIST	=	1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise.

The models above differ most obviously in the controls employed for the uncertainty of clients' operations. Consistent with much prior literature (e.g. Lang and Lundholm 1996; Zhang 2006), Payne (2008) includes forecast dispersion (*DISP*) as a control for the uncertainty in predicting a firm's earnings, while BCK (2008) omit this measure, using the time-series variability of earnings (*STDROE*) instead. Analysts have different private information about the same firms' future earnings, and such differences are reflected in the forecasts they issue. Forecast dispersion reflects analysts' uncertainty about future earnings, and this uncertainty may be caused by the volatility of a firm's underlying fundamentals or by poor quality information (Zhang 2006). Therefore, forecast dispersion may explain greater variability in forecast errors. In fact, the difference in the explanatory power of these controls is stark. I later show that adding a measure of forecast dispersion to BCK's model improves the explanatory power (R^2) by 7.4 per cent.[75] More importantly, a likely reason for BCK's exclusion of *DISP* in their model of forecast accuracy is that their paper also separately tests the impact of audit quality on *DISP*. Whether the possible simultaneous determination of *DISP* and forecast accuracy implies a greater or lesser bias for the test variable than that caused by omitting dispersion is not axiomatically clear. However, given the significant explanatory power of *DISP* when added to BCK's regression, I contend that its inclusion is desirable.

Another significant difference in the models lies in the choice of deflator for the dependent variable and key controls such as dispersion. While BCK deflate forecast errors by stock price, Payne (2008) uses the raw absolute forecast earnings per share

---

[75] The difference in R^2 for BCK models with and without *STDROE* is approximately 0.006 per cent, whereas there is a difference in R^2 of approximately 2 per cent across the Payne model with and without *DISP*.

as the dependent variable. Payne does not explain this choice, but it possibly arises because the likelihood of meeting or beating consensus is modelled separately in Payne's paper, and the undeflated measure is well suited for this purpose. Further, while Payne's control for dispersion is defined consistently with forecast accuracy (it is deflated simply by the mean level of forecast absolute EPS), *ABSECHG* and *PERSIST* are calculated after deflating earnings by beginning stock price. I later show (in Chapter 7) that if one follows the advice of Easton and Summers (2003) and deflates both forecast errors and dispersion by stock price, the explanatory power of Payne's regression model is doubled, and the resulting residuals conform more closely to the assumptions of OLS regression.[76] Additionally, BCK's model includes no control for industry effects, which are likely to be related to both the level of auditor specialisation measured and the accuracy of analysts' forecasts.

A final limitation common to much of the literature and both BCK and Payne's models is the failure to account explicitly for the possible endogenous determination of audit firm industry specialisation. BCK do use a two-stage Heckman approach to model the endogenous determination of audit firm size (Big N), but assume that industry specialisation is exogenously determined.[77] Payne (2008) states that his main results are supported by his untabulated two-stage regressions in which industry specialisation is treated as an endogenous variable. However, the prediction equation (first-stage regression) used is drawn from models of auditor size and includes three 'instruments' (the issuance of stock, the price to earnings ratio and

---

[76] Another common deflator used in the analyst forecasting literature is the firm's absolute actual earnings per share. The untabulated coefficients for industry specialisation are not substantively affected when this deflator is used.

[77] In a sensitivity test, BCK (2008) suggest in the footnote that they also include the IMR as an additional control variable and obtain identical results as to those reported in Models BCK 1a and 1b. However, they do not explicitly indicate whether the IMR is from a regression of Big N audit firm type or industry specialised Big N audit firms.

leverage) that appear to be endogenous to forecast errors. Additionally, these instruments, with the exception of leverage, are of weak explanatory power with respect to audit firm industry specialisation. Identifying exogenous factors with significant explanatory power for the choice of an industry specialist auditor is critically important to two-stage models of endogenous selection (Larker and Rusticus 2010; Lennox et al. 2012).

5.5.1.2 Tests of H1a—Alternative models

To investigate the impact of model specification on the explanation of short-horizon forecast errors, my thesis defines a series of alternative models based on the regressions employed by BCK (2008) and Payne (2008). Models suffixed 'a' measure industry specialisation using a continuous variable, whereas models suffixed 'b' use the dichotomous measure of specialisation. I first define the models that are based on BCK (2008).

BCK's short-horizon regressions:

- BCK original models (Models BCK1a and 1b);
- BCK original model with a control for forecast dispersion (Models BCK 2a and 2b); and
- BCK original model with a control for forecast dispersion, after controlling for the endogenous determination of industry specialist auditors (Models BCK 3a and 3b).

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 HORIZON + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ZSCORE + \beta_6\, LOSS + \beta_7 ABSECHG + \beta_8 STDROE + \beta_9 EL + YEAR + \varepsilon \quad \text{(Model BCK 1a and 1b)}$$

$$ABSFE = \beta_0 + \beta_1 INDSP + \boldsymbol{\beta_2 DISP} + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + \varepsilon$$ (Model BCK 2a and 2b)

Where

ABSFE	=	analysts' absolute forecast errors, measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price;
INDSP	=	the continuous measure of portfolio-share industry specialisation (*INDSP_cont*) as per Equation (8) (all models suffixed 'a') OR
	=	the dichotomous measure of portfolio-share industry specialisation (*INDSP_dum*) as per Equation (9) (all models suffixed 'b');
HORIZON	=	the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date;
SIZE	=	the natural log of the market value of equity;
NUMEST	=	the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting;
ZSCORE	=	Zmijewski's financial distress score, winsorized at 5(-5);
LOSS	=	1 if a firm reports negative earnings, 0 otherwise;
ABSECHG	=	the absolute value of the annual change in EPS, deflated by beginning-of-month stock price;
STDROE	=	the standard deviation of return on equity over the previous five years;
EL	=	earnings per share, winsorized at 5 (-5);
DISP	=	forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by the beginning-of-month stock price; and
YEAR	=	year dummies

The final model is designed to correct for the possible endogenous determination of industry specialisation. Arguably, the factors that affect the choice of auditor may also be related to the difficulty in forecasting earnings, and this effect may not be completely accounted for by the control variables. Thus, factors that determine the choice of industry specialists might be correlated with the residuals from Models BCK 1a, 1b, 2a and 2b. If such endogeneity exists, a two-stage approach (e.g. 2SLS, Heckman treatment regressions or the PSM approach) can potentially mitigate the statistical biases implied. My final short-horizon models, based on BCK (2008), use 2SLS regressions if audit firm industry specialisation is a continuous measure (Model BCK 3a) and PSM regressions if audit firm industry specialisation is measured dichotomously (Model BCK 3b).[78] The first-stage regressions model the level of industry specialisation or the likelihood of appointing a specialist auditor, and they are estimated by OLS regression (for continuous *INDSP*) or a logistic regression (for dichotomous *INDSP*). Below I provide the general form of the first- and second-stage regressions used for 2SLS:

First-stage model (2SLS)

$$INDSP = \beta_0 + \beta_1 INDRELSIZE + \beta_2 CYCLE + \beta_3 DISP + \beta_4 HORIZON + \beta_5 SIZE + \beta_6 NUMEST + \beta_7 ZSCORE + \beta_8 LOSS + \beta_9 ABSECHG + \beta_{10} STDROE + \beta_{11} EL + YEAR + \varepsilon \quad \text{(Model BCK 3a-1)}$$

Where

INDRELSIZE = the sum of the total assets of all firms in an industry at the end of the current year divided by the sum of total assets of all firms at the end of the current year; and

CYCLE = the industry-year adjusted length of operating cycle in days, which is the sum of days' inventory and days' accounts receivable. Days' inventories are estimated using the average of the most current two years' total inventories divided by the sum of cost of goods sold divided by 360. Days' accounts receivable are estimated using the average of the most recent two years of total receivables divided by the sum of sales divided by 360. I adjust this measure for industry effects by subtracting the industry-year mean operating cycle from each observation.

All other variables are defined as above.

[78] I also discuss results generated under Heckman-type endogenous treatment regressions in the sensitivity analyses.

Second-stage model (2SLS)

$$ABSFEadj = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + \varepsilon \quad \text{(Model BCK 3a-2)}$$

Where

ABSFEadj = the industry-adjusted forecast errors, measured as the *ABSFE* minus the industry-year mean of errors.

The first-stage model includes all of the control variables from Models BCK 2a and 2b as well as two instrumental variables, which I argue to be exogenous to forecast errors. First, I include the relative size of an industry (*INDRELSIZE*), because this measure is likely to be strongly positively correlated with the measure of audit firm industry specialisation. The auditor of a firm in an industry comprised of very small firms will likely be measured as having low specialisation even if that auditor has 100 per cent market share in that industry and a small relative share in other industries comprising large firms. I see no theoretical reason that industry relative size should be correlated with forecast accuracy after controlling for firm size and adjusting the forecast accuracy for industry-year effects. My second candidate instrument is clients' operating cycle (*CYCLE*).[79] Francis et al. (1999) argue that firms with longer operating cycles have greater amounts of short-term accruals in working capital accounts. These firms are associated with greater earnings uncertainty and accrual manipulation opportunity, and therefore have stronger incentives to seek a reputable intermediary (such as a high-quality auditor) that can provide additional assurance to signal the reasonableness of reported earnings. Francis et al. (1999) show that firms' operating cycle is positively related to the

[79] As discussed earlier, the use of at least two instrumental variables allows me to conduct over-identification tests to examine the exogeneity of potential instruments.

likelihood of hiring a Big N auditor (a proxy for high-quality auditor), while Godfrey and Hamilton (2005) report a positive association between firms' operating cycle and a high-quality auditor proxied by auditor industry specialisation.

While it is possible that longer operating cycles are associated with riskier earnings (and thus with less accurate forecasts), given that I include several controls for the uncertainty of earnings (i.e. forecast dispersion, absolute earnings change, and variability of earnings) in the second-stage model, it is not obvious that operating cycle will be correlated with the errors in the second-stage models. As noted earlier, valid instruments must be both relevant and exogenous (Larcker and Rusticus 2010). When applied to my data, the LM tests and Wald-type tests found no evidence of under-identification or weak identification (i.e. my instruments are relevant), and the Sargan over-identification tests generated no evidence that the instruments were endogenous to the errors in the second-stage models. Thus, the two standard conditions for instrument validity appear to be satisfied. In developing the two-stage models, I tested several other specifications, which included additional predictors of audit firm industry specialisation as reported by Godfrey and Hamilton (2005). These candidate instruments included leverage, capital intensity, price-earnings ratio, stock issuance and industry regulation, and R&D intensity. While each of these candidate instruments was relevant to the prediction of auditor specialisation, over-identification tests performed on the two-stage models suggested that these variables were not exogenous to forecast accuracy, and thus were not valid instruments.

As the endogenous treatment variable and one of the instruments are a function of industry identity, I do not include industry fixed effects in the two-stage regressions.

To control for industry effects in forecast accuracy, I subtract the industry-year mean of *ABSFE* from the dependent variable for each observation.

In Model BCK 3b, where the dichotomous measure of audit firm industry specialisation is examined, I use the PSM method to correct for the potential endogeneity threats, owing to the advantages of this method over the Heckman treatment regressions, as detailed in Section 5.4.3. Results based on the Heckman regressions are discussed in the sensitivity tests. The models are as below:

First-stage model (PSM)

$$INDSP = \beta_0 + \beta_1 DISP + \beta_2 HORIZON + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ZSCORE + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 STDROE + \beta_9 EL + YEAR + INDUSTRY + \varepsilon \quad \text{(Model BCK 3b-1)}$$

Second-stage model (PSM)

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + INDUSTRY + \varepsilon \quad \text{(Model BCK 3b-2)}$$

I use a logistic regression to estimate the probability of a client selecting an industry specialist. The dependent variable is the dichotomous measure of audit firm industry specialisation, while the independent variables are all control variables used in Model BCK 2b or industry indicator variables. Both Lawrence et al. (2011) and Minutti-Meza (2013) include industry indicator variables in their first-stage matching models and multivariate analysis models. The latter study also shows that the matching procedure with the inclusion of industry and year fixed effects increases the common range of support and the overall fit of the first-stage matching models. As the industry indicator variables are included in both stages, *INDRELSIZE* is largely redundant and is not included in my matching regression.[80] In the second stage, an OLS regression is estimated using the matched sample. To evaluate the impact of post-treatment bias, I re-estimate my PSM models using the one-year lagged value for variables that are likely to respond to audit firm industry specialisation (*DISP*, *ABSECHG*, and *ZSCORE*), or replace these variables with alternates that are theoretically independent of auditor identity (e.g. replace *ABSECHG* with the absolute change in clients' cash flow from operations, *ZSCORE* with leverage, and *EL* with the level of cash flow from operations). These additional tests are discussed in my sensitivity analysis.

I next describe alternate models based on Payne (2008):

- Payne original models (Models Payne 1a and 1b);
- Payne original models with price-deflated forecast errors and price-deflated dispersion (Models Payne 2a and 2b); and
- Payne original models with price-deflated forecast errors and price-deflated dispersion, after controlling for the endogenous determination of industry specialist auditors (Models Payne 3a and 3b).

$$ABSFE\_un = \beta_0 + \beta_1 INDSP + \beta_2 DISP\_eps + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ABSACCR + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 PERSIST + YEAR + INDUSTRY + \varepsilon \quad \text{(Models Payne 1a and 1b)}$$

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ABSACCR + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 PERSIST + YEAR + INDUSTRY + \varepsilon \quad \text{(Models Payne 2a and 2b)}$$

[80] Using *INDRELSIZE* instead of industry indicator variables in the models generates qualitatively similar results, but the first stage logistic regression reports a much smaller Pseudo R^2 (27.5 per cent), relative to 53.21 per cent when industry indicator variables are used.

Where

ABSFE_un	=	analysts' absolute forecast errors (undeflated), measured as the absolute value of the difference between actual *I/B/E/S* earnings per share and forecast earnings per share;
ABSFE	=	analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of the difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price;
INDSP	=	the continuous measure of portfolio-share industry specialisation (*INDSP_cont*) as per Equation (8) (all models suffixed 'a') OR
	=	the dichotomous measure of portfolio-share industry specialisation (*INDSP_dum*) as per Equation (9) (all models suffixed 'b');
DISP_eps	=	forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the absolute value of the mean EPS forecast during the period;
DISP	=	forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price;
SIZE	=	the natural log of total assets;
NUMEST	=	the number of analysts issuing forecasts for the firm in the 90-day window prior to reporting date;
ABSACCR	=	the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets;
LOSS	=	1 if a firm reports negative earnings, 0 otherwise;
ABSECHG	=	the absolute value of the change in annual earnings, deflated by beginning-of-month stock price;
PERSIST	=	1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise;
YEAR	=	year dummies; and
INDUSTRY	=	industry dummies.

The first stages of the 2SLS (PSM) model used in the endogeneity-corrected Payne-type models use identical instruments (matching procedure) to those in the BCK-type models above. Thus, for brevity, I do not formally define them here.

5.5.1.3 Tests of H1b—Audit firm industry specialisation and long-horizon forecast accuracy

To test H1b, I re-estimate each of the Payne- and BCK-type regressions described above using long-horizon forecast accuracy as the dependent variable. When long-horizon forecast accuracy is examined, the identity of a client's auditor in the immediate prior financial year is used in the measurement of audit firm industry specialisation, as this auditor is responsible for the quality of the published financial reports on which long-horizon forecasts are based. Consequently, I use the one-year lag of the instrumental variable *CYCLE* in my long-horizon endogeneity-corrected model (first-stage model under 2SLS). I also test the sensitivity of the long-horizon models to the period over which *DISP*, *ABSECHG*, *ZSCORE*, *ABSACCR* and *PERSIST* are measured, as the current year measures of these variables may be causally affected by audit quality.

5.5.2 Models to Test Hypothesis 2

Hypothesis 2 predicts that the negative relationship between audit firm industry specialisation and forecast errors increases with the client's operating risk. The general form of the model was illustrated in Section 5.2. The selection of controls for testing H2 draws heavily on the model used to test H1b. Both the modified BCK- and Payne-type models capture the most important factors affecting forecast accuracy. However, since the modified BCK model provides greater explanatory power and includes a greater number of controls that are plausibly correlated with my key variables (i.e. firm's operating risk), I report the results based on the modified BCK model.[81] However, I exclude firms' absolute earnings changes (*ABSECHG*), variability in return on equity (*STDROE*) and forecast dispersion (*DISP*) from the model, because cash flow is a component of earnings and volatility in cash flow may mechanically affect volatility in earnings, the inclusion of these earnings uncertainty variables may obscure the interpretation of the results. As noted earlier, *DISP* reflects the degree of disagreement among analysts and other market participants to information uncertainty (Zhang 2006; Imhoff and Lobo 1992; Lang and Lundholm 1996), where the information uncertainty may be highly correlated with the volatility of a firm's underlying fundamentals. Therefore, dispersion is arguably a *response* to firms' cash flow volatility and is excluded from Model (4) below:

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 CFVOL + \beta_3 INDSP*CFVOL + \beta_4 HORIZON + \beta_5 SIZE + \beta_6 NUMEST + \beta_7 ZSCORE + \beta_8 LOSS + \beta_9 EL + YEAR + INDUSTRY + \varepsilon \quad \text{(Model 4)}$$

where

CFVOL = cash flow volatility, measured as the natural log of the 5-year standard deviation of net cash flows from operating activities deflated by average total assets.

I use the dichotomous measure of *INDSP* in testing H2 and employ the PSM method to correct for the endogenous selection of auditors. The dichotomous measure of *INDSP* is employed because it only recognises dominant auditors as having industry specialisation. It is possible that threshold levels of industry portfolio share must be reached before industry expertise accrues (Balsam et al. 2003); thus, the continuous measure of *INDSP* will not capture this feature. In addition, if a continuous measure of *INDSP* is used, a 2SLS model should be employed to correct for endogeneity.

---

[81] The two variables that are included in the Payne's model but not in the BCK's model are absolute accruals (*ABSACCR*) and earnings persistence (*PERSIST*). While the level of earnings, volatility of earnings and absolute earnings change are controlled, the impacts of *ABSACCR* and *PERSIST* on the overall fit of the model are modest. However, when I estimate models including these two variables, the untabulated results are similar to those reported.

However, the 2SLS model might not be easily applied in a model with an interaction term due to the difficulty in finding valid instruments for both the endogenous variable (*INDSP*) and the interaction (*INDSP*CFVOL*) (Baum 2006).[82] Therefore, I use the dichotomous measure of *INDSP* and the PSM method for my tests of H2. To support H2, the coefficient for the interaction term (*INDSP*CFVOL*) must be significant and of negative sign.

5.5.3 Models to Test Hypotheses 3a and 3b

Hypothesis 3a predicts that the negative relationship between audit firm industry specialisation and analysts' absolute forecast errors is decreasing with analyst quality. This is tested by regressing forecast errors against audit firm industry specialisation and its interaction with either each of several analyst quality proxies as per Model 5a, or with the composite score analyst quality proxy as per Model 5b. Following prior studies (Clement 1999; Kim et al. 2011; Drake and Myers 2011), all analyst quality proxies are included in a single regression model because they are potentially correlated with each other. Models 5a and 5b are described below:

$$\begin{aligned} ABSFE = \; & \beta_0 + \beta_1 INDSP + \beta_2 GEXP + \beta_3 INDSP*GEXP + \beta_4 FEXP \\ & + \beta_5 INDSP*FEXP + \beta_6 BSIZE + \beta_7 INDSP*BSIZE + \\ & \beta_8 STAR + \beta_9 INDSP*STAR + \beta_{10} FFOLLOW + \\ & \beta_{11} IFOLLOW + \beta_{12} DISP + \beta_{13} HORIZON + \beta_{14} SIZE + \\ & \beta_{15} NUMEST + \beta_{16} ZSCORE + \beta_{17} LOSS + \beta_{18} ABSECHG \\ & + \beta_{19} STDROE + \beta_{20} EL + INDUSTRY + \varepsilon \end{aligned} \qquad \text{(Model 5a)}$$

$$\begin{aligned} ABSFE = \; & \beta_0 + \beta_1 INDSP + \beta_2 CSCORE1(CSCORE2) + \\ & \beta_3 INDSP*CSCORE1(INDSP*CSCORE2) + \\ & \beta_4 FFOLLOW + \beta_5 IFOLLOW + \beta_6 DISP + \beta_7 HORIZON \\ & + \beta_8 SIZE + \beta_9 NUMEST + \beta_{10} ZSCORE + \beta_{11} LOSS + \\ & \beta_{12} ABSECHG + \beta_{13} STDROE + \beta_{14} EL + INDUSTRY + \varepsilon \end{aligned} \qquad \text{(Model 5b)}$$

[82] Following Baum (2006), I use the interactions of *INDRELSIZE*CFVOL* and *CYCLE*CFVOL* as additional instruments. The Sargan tests of instrument exogeneity fail. I discuss this in more detail in Chapter 8.

Where

GEXP	=	the average general experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where general experience is measured as the number of years through year *t* for which an analyst *i* supplied at least one forecast for any firm;[83]
FEXP	=	the average firm experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where firm experience is measured as number of years through year *t* for which an analyst *i* supplied at least one forecast for firm *j*;
BSIZE	=	the average brokerage size that employs analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where brokerage size is measured as number of analysts employed by a broker employing analyst *i* who follows firm *j* in year *t*;
STAR	=	the proportion of the analysts following firm *j*, during the long-horizon forecast window, who are ranked as an 'All-Star' by *II's* All-America Research Team in year *t;*
CSCORE1 *(CSCORE2)*	=	the composite score, measured as the average of the total ranks for analysts following a firm where the ranking is conducted according to each individual analyst quality proxy within the long-horizon forecast window. *CSCORE1* incorporates the rankings for four proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*), *CSCORE2* incorporates the rankings for three proxies (*GEXP*, *FEXP* and *STAR*);
FFOLLOW	=	the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; and
IFOLLOW	=	the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window.

Tests of H3a use the dichotomous measure of *INDSP* only, for reasons similar to those described in the previous section. H3a is supported if the coefficients for the interactions between analyst quality and audit firm industry specialisation (*INDSP**GEXP*, *INDSP**FEXP*, *INDSP**BSIZE*, *INDSP**STAR*, *INDSP**CSCORE1* and *INDSP**CSCORE2*) are positive and significant.[84]

[83] When the average analyst quality proxies are measured based on all analysts following a firm over the entire year, the untabulated results are similar to the main results.

[84] Similar to the tests of H2, my main regressions for tests of H3a are based on the modified BCK model. However, if I include additional variables from Payne's model (*ABSACCR* and *PERSIST*), the untabulated results are similar to those reported.

H3b regresses the differences in forecast accuracy between the 'worst' analyst and the 'best' analyst based on each of the analyst quality proxies on audit firm industry specialisation and the control variables.

$$\begin{aligned} DIFABSFE = \; & \beta_0 + \beta_1 INDSP + \beta_2 ABSFE\_B + \beta_3 DIFANQ + \\ & \beta_4 ANQ\_B + \beta_5 HORIZON\_B + \beta_6 DIFHORIZON + \\ & \beta_7 DISP + \beta_8 SIZE + \beta_9 NUMEST + \beta_{10} ZSCORE + \\ & \beta_{11} LOSS + \beta_{12} ABSECHG + \beta_{13} STDROE + \beta_{14} EL + \\ & YEAR + INDUSTRY + \varepsilon \end{aligned} \qquad \text{(Model 6)}$$

Where

DIFABSFE	=	the absolute forecast error of the 'worst' quality analyst minus the absolute forecast error of the 'best' quality analyst where the 'worst' and 'best' quality analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
ABSFE_B	=	the absolute forecast error (an inverse function of forecast accuracy) of the 'best' analyst, according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
ANQ_B	=	the level of the quality proxy for the 'best' analyst, where the 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4);*
DIFANQ	=	the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
HORIZON_B	=	the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst, where 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
DIFHORIZON	=	the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4:* and
CSCORE3 (*CSCORE4*)	=	composite score, measured as the aggregate of the analysts' rankings for each individual analyst quality proxy, where the ranking is conducted within the cohort of analysts following a firm in a given year. *CSCORE3* incorporates the rankings for four proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*). *CSCORE 4* incorporates the rankings for three proxies (*GEXP*, *FEXP* and *STAR*).

Once more, I employ the dichotomous measure of *INDSP* in my main tests, and discuss results using the continuous measure of *INDSP* in my sensitivity tests. If audit quality does reduce the difference in forecasting accuracy between the 'worst' and 'best' analysts, as per H3b, the coefficient of *INDSP* is expected to be significantly negative.

5.5.4 Summary of Methods Used to Test Each Hypothesis

In the above sections, I described various estimation methods and discussed the relevance of each method for testing my hypotheses. Table 5.3 summarises the measures of industry specialisation and the regression methods employed in the testing of each hypothesis.

Table 5.3: Summary of Methods

Hypotheses	H1a	H1b	H2	H3a	H3b
INDSP measures					
INDSP (dichotomous)	yes	yes	yes	yes	yes
INDSP (continuous)	yes	yes			sensitivity
Estimation Methods					
OLS (One-way clustering)	yes	yes	yes		yes
OLS (Two-way clustering)				yes	
2SLS	yes	yes			sensitivity
PSM	yes	yes	yes	yes	yes
Heckman	sensitivity	sensitivity	sensitivity	sensitivity	sensitivity

*The cells denoted as 'yes' ('sensitivity') indicate that I use the above measures and estimation methods in the main tests (sensitivity tests).

5.6 Chapter Summary

In this chapter, I discussed the research design used to test the hypotheses developed in Chapter 3. Section 5.2 briefly described the general form of the regression models and the structure of my data. Section 5.3 discussed the various measures for the key independent variables (audit firm industry specialisation, analyst quality and firm's operating risk), dependent variables (analyst forecast accuracy and the difference in forecast accuracy between the 'worst' and 'best' analysts) and control variables. Section 5.4 discussed the estimation methods relevant for each of my hypothesis tests, and Section 5.5 provided the full models. In the following chapter, I describe the data collection process and steps used to determine the final sample for each hypothesis test, followed by a discussion of the descriptive statistics.

CHAPTER 6: SAMPLE AND DESCRIPTIVE STATISTICS

6.1 Introduction

The previous chapter introduced and described the regression models and statistical methods used to test my hypotheses. In this chapter, I describe the data collection process and sample selection for each test (Section 6.2), and report the descriptive statistics and correlations among the variables (Sections 6.3 and 6.4). Section 6.5 concludes the chapter.

6.2 Data Collection and Sample Selection

This section describes the sources of data used in my thesis and the criteria upon which the final samples were determined. The regression models used and the data required for testing Hypothesis 1a, Hypothesis1b and Hypothesis 2 are similar, and thus I describe the data collection and sample derivation for these tests collectively in Section 6.2.1. Tests of Hypothesis 3a and Hypothesis 3b employ analyst quality data, and as such, the final samples and variables differ from those used for testing earlier hypotheses. This is described in Section 6.2.2.

6.2.1 Data Collection and Sample Selection for Tests of H1a, H1b and H2

My sample comprises U.S. firms for the period 1989 to 2010. The sample begins in year 1989 because this is the earliest year common to the papers most closely related to my study (BCK 2008; Payne 2008).[85] Payne's (2008) sample begins in 1989,

---

[85] I re-estimate the tests on the relationship between analyst forecast accuracy and audit firm industry specialisation by restricting the sample to (a) the years included in BCK's study (1996–2001) and (b) Payne's study period (1989–2005). I obtain results qualitatively similar to my main results.

likely because this is the year in which the Big 8 audit firms became the Big 6. Analysts' forecasts and their respective actual values were obtained from the *I/B/E/S* Detail History file. Client firms' financial data and the data used to calculate audit firm industry specialisation were obtained from *COMPUSTAT* North America Industrial file.[86]

As the majority of my empirical tests examine long-horizon forecast accuracy, I focus on describing the derivation of the long-horizon forecast accuracy sample. I applied similar procedures to derive the short-horizon sample; however, for brevity, I do not discuss the derivation of this sample at length here (see Appendix C for details). Table 6.1 describes the derivation of the sample used to test H1b and H2. The sample derivation for tests using the Payne-type models (BCK-type models) is illustrated in the first (last) two columns of Table 6.1. There are 62,385 firm-year observations for which there are sufficient data to compute long-horizon consensus forecast errors, of which 13,791 (24,757) firm-year observations are missing financial information on *COMPUSTAT* necessary to estimate the Payne-type models (BCK-type models). The greater sample attrition for the BCK-type models is due to the additional data required to estimate the five-year standard deviation of return of equity (*STDROE*), and Zmijewski's financial distress score (*ZSCORE)* (a large number of firms in the financial sector do not have the current assets and current liabilities that are required to compute the liquidity ratio for estimating *ZSCORE*).

---

[86] In my sensitivity analysis, I use audit fees (rather than clients' assets) to calculate audit firm industry specialisation. Audit fees were obtained from Audit Analytics for the period 2000 to 2010.

I now focus my description on the sample derivation using Payne-type models.[87] Following Payne (2008), I restrict my sample to clients of Big N firms to control for any potential factors that might affect my analysis induced by audit firm size. This reduces my sample by 5,278 observations. Financial sector firms (firms with SIC codes between 6000 and 6999) are then excluded (4,689 firm-year observations)[88] due to the special regulatory requirements imposed on this industry, which may affect firm's choice of industry specialist audit firms and cause abnormal accrual estimation. Following Payne (2008), I require a minimum of 20 observations in each two-digit SIC industry-year to obtain a stable indicator of audit firm industry specialisation, reducing the sample by a further 551 observations. Thirteen firm-year observations are excluded, as these firms received modified audit opinions, which would likely mitigate the potential impact of modified audit opinions on the quality of financial statements. To enable the meaningful calculation of price-deflated analyst forecast dispersion (*DISP*), I require at least three analysts to issue forecasts within a given window and for one-month lagged stock prices to be available. This reduces my sample by a further 9,249 observations. Finally, following Payne, I exclude 1,530 extreme observations (i.e. $1^{st}/99^{th}$ percentile of each untransformed continuous variable) to minimise the impact of outliers on the regression results.[89] Therefore, the final sample consists of 27,284 firm-year observations for the long-horizon sample using regressions based on Payne (2008). The additional data requirements of the endogeneity-corrected regressions (the one-year lag of client

---

[87] BCK (2008) do not detail their sample construction procedure, whereas Payne (2008) explicitly shows the impact of each procedure on the sample. Thus, I largely follow Payne's procedures to derive my sample.

[88] The data requirement for non-financial firms only reduces BCK's sample by 725 firm-year observations because many financial firms are dropped in the first stage when I require data to compute *ZSCORE*.

[89] To be consistent with Payne and BCK's approaches, the control variables, firm size (*SIZE*) and analyst following (*NUMEST*) are not truncated because the natural log values of the variables are used. Earnings per share (*EL*) is winsorized at 5 (-5).

firm's operating cycles, absolute accruals, absolute earnings change, earnings persistence and analyst forecast dispersion) further reduce the sample employed in those models (21,232 firm-years). A similar sample construction procedure applied to short-horizon forecast errors generates a sample of 31,806 firm-year observations (31,358 for endogeneity-corrected models). Despite my study spanning an additional five years, my sample is just 3,670 observations greater than Payne's. There are two likely reasons for this. First, Payne uses the *I/B/E/S* Summary File, which contains a number of 'stale' forecasts, and which may therefore result in fewer observations being filtered due to low analyst coverage. Second, following the 2008 Global Financial Crisis, *I/B/E/S* retrospectively deleted all forecasts issued by a number of brokers.[90] The equivalent samples employed in the BCK-type regressions are slightly smaller (25,489 in the single-stage models and 21,119 in the two-stage models).[91]

Hypothesis 2 is tested by augmenting the BCK-type models with main effects and interaction terms incorporating cash flow volatility. The availability of data necessary to estimate this additional measure reduces my sample to 23,558 cases for testing H2 (in Panel B of Table 6.1).

---

[90] For example, in the 2010 detail tape update, *I/B/E/S* permanently removed forecasts made by 16 brokers (WRDS 2010 *I/B/E/S* notes).

[91] Without the inclusion of the standard deviation of return on equity (*STDROE*), my sample for the single-stage BCK-type regressions contains 27,904 firm-year observations. This is similar to the 27,284 derived under the Payne-type models.

Table 6.1: Sample Selection for Tests of H1b and H2

Panel A Tests of H1b

	Payne-type models		BCK-type models	
Available **long-horizon** consensus forecasts		62,385		62,385
less: unavailable financial information from *COMPUSTAT*	−13,791	48,594	−24,757	37,628
less: non-Big N firms	−5,278	43,316	−3,277	34,351
less: financial sector firms	−4,689	38,627	−725	33,626
less: firms in industries with less than 20 members in a given year	−551	38,076	−491	33,135
less: firms subject to modified audit opinions	−13	38,063	−11	33,124
less: observations where forecast dispersion cannot be calculated	−9,249	28,814	−6,903	26,221
less: extreme observations for continuous variable (1/99 percentile)	−1,530		−732	
Final sample of long-horizon forecast (used in single-stage models)		**27,284**		**25,489**
less: additional data requirements for the endogeneity-corrected regressions	−6,052		−4,370	
Final sample of long-horizon forecast (used in 2SLS)		**21,232**		**21,119**

Panel B: Tests of H2

Available sample from H1b		25,489
less: insufficient data to estimate cash flow volatility	−1,931	
Final sample for H2		**23,558**

6.2.2 Data Collection and Sample Selection for Tests of H3a and H3b

The samples used to test H3a and H3b begin in year 1993 because this is the earliest year for which I have access to data regarding the identity of 'All-Star' analysts. I obtain data to compute analyst quality proxies (except for the 'All-Star' proxy) from the *I/B/E/S* detail file. To compute the average value of analyst quality proxies at the firm-year level, I first need to calculate individual analysts' general experience, firm-specific experience and their employer size. I identify the individual analysts and brokers either by reference to the *I/B/E/S* broker translation file (which has not been updated since 2006 and is no longer publicly distributed), or by cross-referencing to the analyst and broker names and codes reported in the *I/B/E/S* Detail Recommendations file. To track analyst experience over as long a period as possible, the sample period used to compute the experience proxies spans 1983–2010, as the *I/B/E/S* database identifies most contributing analysts from 1983.[92] This is consistent with the approach taken in prior studies (Emery and Li 2009; Clement 1999). Clement (1999, 294) acknowledges that the *I/B/E/S* data set is left censored such that an analyst experience prior to the first year of available data is unknown. In particular, year dummy variables are likely to be systematically correlated with the estimated experience measures, even though they may be exogenous to analysts' (unobservable) true experience. While Clement (1999) estimates experience using data from 1983–1994, he conducts regression analysis for the period 1985–1994 to mitigate the left censoring problem. This approach is widely adopted in later studies (Brown 2014; Emery and Li 2009; Lee 2004). The 10-year pre-sample window (1983–1992) in my study reduces, but does not eliminate, mechanical bias arising from the left censoring bias. To address the potential collinearity between year dummies and the experience variables, I estimate models using the OLS with standard errors, clustering by firm and year rather than including year dummies.

I identified 'All-Star' analysts by reference to the rankings published in the back archive of the *II* magazine. 'All-Star' rankings for the years 1993–1999 with matched analyst and broker codes were kindly provided by Craig Brown, and were manually collected from the magazine for the remaining sample years (2000–2010). This source provides analysts' full names, their employer firm, their rankings (1st

[92] For the years prior to 1983, the *I/B/E/S* Detail History file includes 230 identifiable analysts (i.e. analyst codes that are non-zero and non-missing) who followed 108 firms. For 1983, there are 2,280 identifiable analysts following 2,745 firms. Coverage increases thereafter.

place, 2nd place, 3rd place or runner-up) [93] and the industry sector in which these analysts were awarded their rankings. Consistent with much prior literature (Loh and Stulz 2010; Marcus et al. 2014; Chan et al. 2004), I recognised an analyst as an 'All-Star' if the analyst is ranked (including runner-ups) in a particular year, regardless of the industry in which a particular client firm is a member.[94]

Table 6.2 Panel A describes the procedures used in constructing my sample for tests of H3a. I eliminate 2,710 firm-year observations from the available sample (25,489) used in the tests of H1b due to the missing data for the 'All-Star' rankings prior to 1993. Following Jacob et al. (1999) and Clement (1999), if analyst and broker codes are missing or equal to '0', or analyst names do not indicate an individual analyst but correspond to an industry or team grouping (i.e. the names provided include company, group or industry names), these analysts are excluded from the analyst quality sample. Accordingly, firm-year observations where an insufficient number of analysts/brokers can be clearly identified according to this criterion are eliminated (31). My final sample for tests of H3a consists of 22,742 observations.

[93] Analysts' rankings are determined by numerical score. An analyst is designated runner-up when the analyst's score falls within 35 per cent of third-analyst's score (*II's* magazine).
[94] Some prior studies (Emery and Li 2009) classify analysts as an 'All-Star' only in the industries in which they are ranked. Others (Chan et al. 2004; Marcus et al. 2014) do not differentiate according to industry classification. To preserve my 'All-Star' sample size in testing H3b (where I require a firm to be followed by at least one 'All-Star' analyst), I do not differentiate the industry classification. Since prior studies (Stickel 1992; Chan et al. 2004) show that 'All-Star' analysts' superior forecast accuracy relative to other analysts persists after the selection year, I conduct additional analysis, which includes the testing of multiple variations in the time window over which an analyst is identified as an 'All-Star'.

Table 6.2: Sample Selection for Tests of H3a and H3b

Panel A Tests for H3a

Available sample from H1b		25,489
less: firm-year observations prior to 1993 due to the missing value for the 'All-Star' rankings	−2,716	22,773
less: firm-year observations where mean analyst quality proxies cannot be computed (i.e. firm-year observations followed exclusively by analysts whose analyst codes on *I/B/E/S* are missing or analyst codes equal to '0'; or whose analyst names include 'industry', 'group' or 'company' name)	−31	
Final sample for tests of H3a		**22,742**

Panel B Tests of H3b (Experience, Brokerage Size and Composite Score Proxies)

Available sample from H3a		22,742
less: extreme observations of the dependent variable (1/99 percentile)	−601	22,141
less: cases where the 'worst' and 'best' analysts have equivalent measures of a quality proxy	−1,616	
Final sample for tests of H3b using the experience, brokerage size and composite score proxies		**20,525**

Panel C Tests of H3b ('All-Star' Proxy)

Available sample from H3a		22,742
less: extreme observations of the dependent variable	−251	22,491
Less: firms that do not have at least one 'All-Star' and one non-star analyst	−13,215	
Final sample for tests of H3b using the 'All-Star' proxy		**9,266**

Table 6.2 Panel B describes the procedures used in deriving my sample for tests of H3b. The dependent variable in tests of H3b is the difference in absolute forecast errors of the 'worst' and 'best' analysts (where the 'worst' and 'best' analysts are determined for each of the analyst quality proxies). I first exclude extreme observations for the dependent variables,[95] reducing my sample to 22,141. I then eliminate 1,616 cases where a firm's 'worst' and 'best' quality analysts have identical values for a quality proxy (e.g. analysts who have the same length of general or firm-specific experience; or analysts whose employers' sizes are the

[95] The truncation is conducted after computing the difference in absolute forecast errors of 'worst' and 'best' analysts.

same).[96] Therefore, when I test H3b using experience, brokerage size and composite score proxies for analyst quality, my final sample is reduced to 20,525 firm-year observations. Tests using the 'All-Star' proxy are further constrained by the requirement that a firm is followed by at least one analyst on the 'All-Star' list and one analyst not on the 'All-Star' list. This requirement and the exclusion of extreme cases give a final sample of 9,266 for tests of H3b, which use 'All-Star' status as the proxy for analyst quality.

6.3 Descriptive Statistics

The descriptive statistics for the variables included in the models are presented in Section 6.3.1. Section 6.3.2 reports the univariate comparisons of the means and medians of the variables across two groups—clients of industry specialists and non-specialists—for the unrestricted samples. Section 6.3.3 shows the differences in the mean value of the covariates across these two groups in the propensity score matched samples.

6.3.1 Descriptive Statistics

Table 6.3 presents the descriptive statistics for the firm-year-level measures of analyst-related variables (Panel A), auditor and client firm variables used in all tests (Panel B), and the additional variables employed for tests of H3b (Panel C).

Panel A tabulates the analyst-related variables measured at firm-year level. The tabulated statistics for analyst short-horizon absolute forecast errors, forecast dispersion and forecast horizon variables (rows 1 to 6) are based on the short-horizon

---

[96] Without this sample restriction, the untabulated regressions show results that are similar to my main results.

endogeneity-corrected BCK models used to test H1a. The descriptive statistics reported for long-horizon absolute forecast errors (rows 7 to 12) and other analyst-related variables (rows 13 to 18) are based on the long-horizon sample used for tests of H3a.[97] Unsurprisingly, analysts' earnings forecasts are more accurate at year-end than at the beginning of the year. The mean (median) price-scaled short-horizon forecast errors (*ABSFE*) are 0.012 (0.002), relative to mean (median) price-scaled long-horizon forecast errors of 0.026 (0.008). While the mean of the short-horizon forecast errors are slightly smaller than that reported by BCK (2008), their median value is very close to BCK's (0.003). Similarly, the mean and median price-scaled forecast dispersion (*DISP*) are smaller at year-end than for those forecasts made at the beginning of the year. This increased consensus in earnings reflects the change in available information at year-end. The distributions of the analyst quality and analysts' portfolio complexity variables are similar to those reported by Drake and Myers (2011), who cover a shorter sample period (i.e. 1993–2008). For example, analysts in my sample have an average general experience (*GEXP*) of 6.938 years and an average firm-specific experience (*FEXP*) of 3.266 years, which is comparable to the 7.1 years and 3.5 years, respectively, tabulated by Drake and Myers (2011). On average, a firm is followed by analysts whose brokerage firm (*BSIZE*) employs approximately 56 analysts; and 10.9 per cent of the analysts covering a firm have been ranked as 'All-Star' analysts (*STAR*) by the *II* ranking system. The portfolio complexity variables show that on average an analyst follows approximately 17 firms (*FFOLLOW*) and five industries (*IFOLLOW*). This is consistent with the statistics reported by Drake and Myer (2011).

[97] The untabulated descriptive statistics for long-horizon analyst-related variables in the tests of H1b and H2 are similar to those reported in Table 6.3.

Panel B reports the distributions of my auditor and client financial data. Approximately 38 per cent of the firms are clients of an industry specialist audit firm, which is slightly higher than the 35.15 per cent reported by Payne for an overlapping, but earlier, sample period. Firm size (*SIZE*), as measured by mean assets (market capitalisation) is approximately $5.320 billion ($6.065 billion). The firms included in my sample are slightly larger compared to those included in Payne and BCK's studies, reflecting the relative recency of my sample. An average of seven analysts (*NUMEST*) follow a firm in a given year, which is slightly lower than the nine reported by Payne (2008) in examining the short-horizon forecasts. The lower number of analyst following in my long-horizon sample reflects analysts' tendency to issue forecasts more frequently at year-end. Firms in my sample have similar levels of uncertainty/risk (*ZSCORE* and *ABSACCR*) to those in BCK and Payne's samples. Approximately 15.8 per cent of the firm-year observations experience a loss (*LOSS*) in the current year, compared to 14.4 per cent in Payne's sample.[98] Variables capturing earnings variability (*ABSECHG*, *STDROE*, *PERSIST*) have distributions similar to those reported by Payne (2008) and BCK (2008), with the exception of earnings per share (*EL*), which has a mean of 1.043 dollars compared to the 42.6 cents reported by BCK (2008). This difference in *EL* may be attributable to my use of the *I/B/E/S*-defined EPS rather than the GAAP operating profit from *COMPUSTAT*. The average earnings per share (*EL*) calculated based on the GAAP operating profit is 79 cents. The use of *I/B/E/S*-defined earnings appears preferable, as these reflect the 'Pro Forma' earnings that are the focus of analysts' forecasts, upon which my dependent variables are based. My primary operating risk proxy, client cash flow volatility (*CFVOL*), has an average value of 0.05, which is similar to

[98] If *LOSS* is computed using the net earnings available on *COMPUSTAT*, the mean is 0.24, compared to 0.35 reported by BCK (2008).

that reported by Dichev and Tang (2009) (0.039) over a shorter sample period (1988–2004).

Panel C shows the descriptive statistics for the difference in absolute forecast errors of the 'worst' and 'best' analysts for each of the analyst quality proxies and additional control variables employed for testing H3b. As expected, the average absolute forecast errors of the 'worst' analysts (*ABSFE_W*) are slightly larger than or equal to those of the 'best' analysts (*ABSFE_B*). On average, the 'best' and 'worst' analysts covering a firm have an approximately eleven-year difference in general experience and a six-year difference in firm-specific experience (*DIFANQ*). The mean difference in the number of analysts employed by a brokerage firm is 6.343. The 'best' analysts covering a given firm on average have around 13.526 years' (7.174 years') general experience (firm-specific experience) and are employed by a brokerage firm with around 118 analysts (*ANQ_B*). There is not a large difference in the mean value of forecast timing between the 'worst' and 'best' analysts (*DIFHORIZON*). However, it is notable that the standard deviation of the *DIFHORIZON* is relatively large (ranging from 20–32 days), which implies that the variations in forecast timing may have an impact on these analysts' forecasting performance.

Table 6.3: Descriptive Statistics
Panel A Analyst-Related Variables Used in the Tests of H1a–H3a

	MEAN	SD	Q1	MEDIAN	Q3
			Short-horizon (n=28,701)		
*ABSFE_un**	0.093	0.165	0.013	0.038	0.098
ABSFE	0.012	0.07	0.001	0.002	0.007
*DISP_eps**	0.118	0.275	0.014	0.035	0.094
DISP	0.006	0.013	0.001	0.002	0.005
HORIZON (Log)	3.199	0.782	2.89	3.346	3.717
HORIZON(Days)	25	2	18	28	41
			Long–horizon (n=22,742)		
*ABSFE_un**	0.329	0.485	0.057	0.156	0.389
ABSFE	0.026	0.053	0.003	0.008	0.024
*DISP_eps**	0.177	0.402	0.025	0.061	0.152
DISP	0.009	0.017	0.001	0.003	0.009
HORIZON(Log)	5.712	0.100	5.677	5.709	5.745
HORIZON(Days)	304.470	34.853	285.000	296.000	316.000
GEXP	6.938	2.752	5.000	6.778	8.500
FEXP	3.266	2.006	1.800	2.857	4.333
BSIZE	56.477	29.395	34.800	53.400	73.615
STAR	0.109	0.166	0.000	0.000	0.191
FFOLLOW	16.860	6.988	13.250	16.000	19.000
IFOLLOW	5.213	2.636	3.286	4.750	6.500

**The statistic for this variable is computed based on the sample used for the endogeneity–corrected Payne's model.*

Panel B Audit Firm Industry Specialisation and Firm–Level Variables Used in the Tests of H1a–H3a

	MEAN	SD	Q1	MEDIAN	Q3
INDSP_dum	0.381	0.486	0.000	0.000	1.000
INDSP_cont	0.049	0.044	0.016	0.043	0.069
SIZE (asset)($B)	5.320	16.483	0.322	0.994	3.473
SIZE (market cap) ($B)	6.065	19.664	0.400	1.144	3.662
NUMEST(number)	6.806	5.489	3.000	5.000	9.000
NUMEST(log)	1.653	0.714	1.099	1.609	2.197
ZSCORE	–1.619	1.536	–2.668	–1.655	–0.756
*ABSACCR**	0.083	0.071	0.036	0.066	0.110
LOSS	0.158	0.365	0.000	0.000	0.000
*PERSIST**	0.600	0.480	0.000	1.000	1.000
ABSECHG	0.036	0.063	0.007	0.014	0.036
STDROE	0.074	0.101	0.019	0.039	0.083
EL	1.043	1.272	0.280	0.870	1.680
*CFVOL (Raw)***	0.050	0.043	0.023	0.038	0.062
*CFVOL (Log)***	–3.270	0.740	–3.750	–3.260	–2.780

**The statistic for this variable is computed based on the sample used for the endogeneity-corrected Payne's model.*
***The statistic for this variable is computed based on the sample for tests of H2.*

Panel C Additional Variables Used in the Tests of H3b

	DIFABSFE	*ABSFE_W*	*ABSFE_B*	*DIFANQ*	*ANQ_B*	*HORIZON_B* (days)	*DIFHORIZON* (days)
Difference Based on *GEXP* (n=20,525)							
MEAN	0.000	0.024	0.024	11.296	13.526	304.240	−1.143
MEDIAN	0.000	0.009	0.009	11.000	13.000	296.000	0.000
SD	0.014	0.047	0.047	6.064	5.676	34.638	30.782
Difference Based on *FEXP* (n=20,525)							
MEAN	0.000	0.024	0.024	6.343	7.174	303.873	−1.594
EDIAN	0.000	0.009	0.009	5.000	6.000	296.000	0.000
SD	0.013	0.047	0.047	4.855	4.845	33.966	29.020
Difference Based on *BSIZE* (n=20,525)							
MEAN	0.000	0.024	0.024	6.343	117.930	302.560	3.555
MEDIAN	0.000	0.009	0.008	5.000	127.000	294.000	0.000
SD	0.015	0.047	0.046	4.855	58.050	33.986	33.161
Difference Based on *STAR* (n=9,266)							
MEAN	0.000	0.020	0.020				2.920
MEDIAN	0.000	0.008	0.008				3.750
SD	0.007	0.041	0.041				20.781
Difference Based on *CSCORE3* (n=20,525)							
MEAN	0.000	0.024	0.024	13.942	23.475	302.966	1.209
MEDIAN	0.000	0.008	0.009	8.500	17.000	295.000	0.000
SD	0.015	0.047	0.047	14.920	18.811	34.345	31.989
Difference Based on *CSCORE4* (n=20,525)							
MEAN	0.000	0.024	0.024	11.396	18.062	303.518	−0.558
MEDIAN	0.000	0.009	0.009	7.500	13.000	295.000	0.000
SD	0.014	0.047	0.047	11.510	14.311	34.493	30.763

This table presents the descriptive statistics for firm-year-level measures of analyst-related variables (Panel A), auditor and client firm variables used in all tests (Panel B) and the additional variables employed in the tests of H3b (Panel C). Variable definitions are provided in Table 6.7.

6.3.2 Univariate Comparisons (Full Sample)

Table 6.4 reports the univariate comparisons of variables across clients of industry specialist auditors and those of other auditors for my full samples. Panel A presents the differences in means and medians across the variables used in testing H1a, H1b, H2 and H3a. Clients of industry specialists are generally larger (*SIZE*), have greater analyst following (*NUMEST*), experience less severe financial distress (*ZSCORE*), have a higher level of accruals (*ABSACCR*) and likelihood to incur financial losses (*LOSS*), have less persist earnings (*PERSIST*), lower earnings per share (*EL*) and more volatile returns on equity (*STDROE*).[99] These statistics are generally consistent with those reported by Payne (2008). In summary, all of the firm-level control variables significantly differ in means between clients of industry specialist audit firms and those of non-specialists, emphasising the importance of conducting additional analysis on propensity score matched samples. Panel B reports the differences in means and medians across the additional controls included in tests of H3b. I find significant mean and/or median differences in every case except for the variables relating to forecast horizon.

[99] The t-statistics are reported based on the differences in the mean values of the variables, where differences equal the mean value of a non-specialist's client characteristics minus that of a specialist's client characteristics. Therefore, a negative t-statistic indicates a greater mean value of a specialist compared to a non-specialist.

Table 6.4: Univariate Comparisons of Regression Variables (Full Sample n=22,742)

Panel A Variables Used in the Tests of H1a–H3a

	(1) Industry Specialist (38.12%)			(2) Non-industry Specialist (61.88%)			Differences (2) - (1)	
	MEAN	MEDIAN	SD	MEAN	MEDIAN	SD	t-stat	wilcoxon
ABSFE (SH)	0.025	0.008	0.05	0.027	0.009	0.056	−0.85	−3.146***
DISP (SH)	0.009	0.004	0.016	0.008	0.003	0.015	−3.43***	−6.759***
ABSFE (LH)	0.025	0.008	0.052	0.026	0.009	0.054	3.9***	2.321**
DISP(LH)	0.01	0.004	0.018	0.008	0.003	0.015	−7.59***	−7.355***
HORIZON (SH)	3.23	3.36	0.755	3.18	3.34	0.796	−3.5***	−4.051***
HORIZON (LH)	5.714	5.711	0.117	5.711	5.707	0.089	−2.46**	−4.702***
GEXP	6.666	6.417	2.769	7.106	7	2.728	11.75***	13.972***
FEXP	3.097	2.625	1.969	3.37	3	2.021	9.97***	11.435***
BSIZE	56.459	53.333	28.794	56.488	53.5	29.76	0.07	−0.291
STAR	0.096	0	0.154	0.117	0	0.172	9.45***	8.994***
FFOLLOW	16.771	15.667	7.632	16.915	16	6.56	1.51	7.035***
IFOLLOW	5.055	4.733	2.348	5.311	4.8	2.794	7.12***	3.195**
SIZE (asset) ($B)	6.375	0.963	18.486	4.669	1.01	15.081	−7.59***	2.546***
SIZE (market cap) ($B)	8.031	1.262	24.911	5.084	0.994	18.422	−4.29***	−2.832***
NUMEST	1.675	1.609	0.73	1.627	1.609	0.703	−4.95***	−4.186***
ZSCORE	−1.539	−1.646	1.761	−1.668	−1.659	1.377	−6.16***	−1.851***
*ABSACCR**	0.092	0.069	0.079	0.082	0.065	0.069	−10.28***	−8.251***
LOSS	0.229	0	0.42	0.114	0	0.317	−23.56***	−23.281***
*PERSIST**	0.591	1	0.492	0.609	1	0.488	3***	2.995***
ABSECHG	0.037	0.014	0.065	0.036	0.014	0.062	−1.72*	0.045
STDROE	0.098	0.051	0.126	0.059	0.033	0.078	−28.93***	−25.217***
EL	0.825	0.682	1.312	1.178	0.98	1.227	20.53***	21.17***
*CFOVOL***	0.056	0.04	0.051	0.047	0.037	0.037	−15.42***	−8.243***

**The statistic for this variable is computed based on the sample used for the endogeneity-corrected Payne's model.*
***The statistic for this variable is computed based on the sample for tests of H2.*

Panel B Additional Variables Used in the Tests of H3b (Based on *CSCORE4* Analyst Quality Proxy n=20,525)

	MEAN	MEDIAN	SD	MEAN	MEDIAN	SD	t-stat	wilcoxon
DIFABSFE	0	0	0.015	0	0	0.015	2.23**	1.465***
ABSFE_W	0.025	0.009	0.048	0.025	0.009	0.048	2.31**	2.623**
ABSFE_B	0.024	0.009	0.048	0.024	0.009	0.048	1.61	2.174**
DIFANQ	13.579	8.5	14.324	13.579	8.5	14.324	−4.46***	−3.519***
ANQ_B	24.387	17.5	20.087	22.918	17	17.965	−5.43***	−3.867***
HORIZON_B (days)	303.826	295	41.05	302.441	295	29.499	−2.8	−0.605
DIFHORIZON (days)	0.928	0	31.301	0.928	0	31.301	−1.61	−0.094

This table presents the univariate comparisons of variables across clients of industry specialist auditors and those of other auditors for my full samples. Panel A presents the differences in means and medians across the variables used in the tests of H1a, H1b, H2 and H3a. Panel B reports the differences in the means and medians across the additional controls included in the tests of H3b. Variable definitions are provided in Table 6.7.

6.3.3 Univariate Comparisons (Propensity Score Matched Sample)

I address the potential endogenous selection of specialist auditors using a variety of approaches, including the use of a propensity score matched sample. This approach produces matched sample based on propensity scores generated from a logistic regression of auditor selection against factors posited to affect this choice. Table 6.5 reports the mean differences across clients of industry specialists and non-specialists for all control variables, thereby testing the sample-balancing effectiveness of the propensity matching approach. For brevity, I only tabulate the statistics for the matching conducted in the tests of H1b, and the tests of H3a and H3b when the composite score measure (constructed based on four analyst quality proxies) is used. Using the one-to-one nearest neighbour within-caliper matching approach, my matched samples represent between 16 and 20 per cent of the full samples. The final calipers used range between 1 and 2.5 per cent (i.e. the treatment and control firms have an absolute difference in propensity scores of less than 1 or 2.5 per cent). Importantly, there are no significant differences in the means of any covariates across the two groups.

I also evaluate the range of common support and the predictive power of each of my matching models. I plot histograms (untabulated) for the propensity scores of the treatment firms (clients audited by an industry specialist) and control firms (clients audited by a non-specialist) after each matching regression, and observe that control firms' propensity scores generally span the full range of the propensity scores of the treatment firms. I tabulate the first stage of the PSM regressions in Appendix D. The pseudo R^2 statistics are generally above 48 per cent and the area under the ROC for all first-stage matching models is above 90 per cent, which suggests that my matching models are of great predictive power. Among the predictors, client size (*SIZE*) and volatility in return on equity (*STDROE*) appear to be the most important factors explaining the choice of industry specialist auditors, as these variables are correlated with *INDSP* across all models where these variables are included.

In summary, the PSM approach adopted is of great predictive power for the selection of auditors and has been successful in generating a balanced sample, at least in terms of the determinants of auditor selection identified in my thesis.

Table 6.5: Univariate Comparisons of Means (Propensity Score Matched Samples)

Panel A Tests of H1b (BCK-type Models)

	Industry Specialist Mean (1)	*Non-industry Specialist* Mean (2)	*p*-values
DISP	0.009	0.009	0.573
HORIZON	5.709	5.711	0.441
SIZE	6.782	6.822	0.429
NUMEST	1.668	1.676	0.711
ZSCORE	−1.913	−1.897	0.701
LOSS	0.145	0.153	0.42
ABSECHG	0.039	0.039	0.938
STDROE	0.072	0.077	0.1
EL	0.918	0.903	0.654
N	2,421	2,421	

First-stage model: *INDSP = DISP + HORIZON + SIZE + NUMEST + ZSCORE + LOSS + ABSECHG + STDROE + EL + YEAR + INDUSTRY*
Matching caliper size 0.0125

Panel B Tests of H1b (Payne-type Models)

	Industry Specialist Mean (1)	*Non-industry Specialist* Mean (2)	*p*-values
DISP	0.009	0.009	0.655
SIZE	6.722	6.796	0.122
NUMEST	10.408	10.635	0.311
ABSACCR	0.095	0.092	0.273
LOSS	0.142	0.141	0.875
ABSECHG	0.039	0.038	0.598
PERSIST	0.603	0.601	0.866
N	2,649	2,649	

First-stage model: *INDSP = DISP + SIZE + NUMEST + ABSACCR + LOSS + ABSECHG + PERSIST + YEAR + INDUSTRY*
Matching caliper size 0.025

Panel C Tests of H3a

	Industry Specialist Mean (1)	*Non-industry Specialist* Mean (2)	*p*-values
FFOLLOW	17.739	17.742	0.992
IFOLLOW	5.627	5.599	0.746
DISP	0.010	0.010	0.340
HORIZON	5.710	5.708	0.507
SIZE	6.950	6.923	0.641
NUMEST	1.694	1.704	0.673
ZSCORE	−1.852	−1.875	0.652
LOSS	0.154	0.167	0.302
ABSECHG	0.040	0.038	0.505
STDROE	0.076	0.079	0.239
EL	0.945	0.948	0.941
N	1,843	1,843	

First-stage model: *INDSP = FFOLLOW + IFOLLOW + DISP + HORIZON + SIZE + NUMEST + ZSCORE + LOSS + ABSECHG + STDROE + EL + YEAR + INDUSTRY*
Matching caliper size 0.01

Panel D Tests of H3b

	Industry Specialist Mean (1)	*Non-industry Specialist* Mean (2)	*p*-values
ABSFE_B	0.0250	0.027	0.378
DIFANQ	15.593	16.470	0.157
ANQ_B	9.995	10.282	0.113
DISP	0.008	0.008	0.972
HORIZON_B	302.920	303.350	0.692
DIFHORIZON	0.718	1.826	0.324
SIZE	7.071	7.056	0.805
NUMEST	1.783	1.806	0.366
ZSCORE	−1.854	−1.883	0.562
LOSS	0.135	0.146	0.343
ABSECHG	0.036	0.035	0.544
STDROE	0.075	0.079	0.167
EL	1.011	1.0145	0.941
N	1,662	1,662	

First-stage model: *INDSP = DIFANQ + ANQ_B + DISP + HORIZON_B + DIFHORIZON + SIZE + NUMEST + ZSCORE + LOSS + ABSECHG + STDROE + EL + YEAR + INDUSTRY*
Matching caliper size 0.025

This table presents the mean differences across clients of industry specialist auditors and non-specialists for all control variables for tests of H1b using Payne-type models (Panel A) or BCK-type models (Panel B), tests of H3a (Panel C) and tests of H3b (Panel D). Variable definitions are provided in Table 6.7.

6.4 Correlation Analysis

This section reports the analyses in which the potential multicollinearity issues are identified and addressed. Collinearity refers to a strong linear relationship between two independent variables. Multicollinearity occurs when two or more independent variables in a regression model are highly correlated with one another. A high degree of multicollinearity in multiple regression analysis may inflate the standard errors of the coefficients even though the R^2 for the regression is high, and may cause the parameter estimates to be highly sensitive to small changes in the data (Greene 2002, 57). Researchers commonly calculate the Variance Inflation Factor (VIF) to detect the occurrence of multicollinearity (O'Brien 2007). VIFs show to what extent the estimated variance of a regression coefficient exceeds the level that would be obtained if the independent variables were uncorrelated with other variables in the model (O'Brien 2007, 674). As a general rule in the literature, values of VIF greater than 10 may be taken as an indication for multicollinearity concern (Neter et al. 1989, 409; Kennedy 1992, 183; Marquardt 1970).

Table 6.6 provides the correlation matrix for the key variables, including long-horizon absolute forecast errors, difference in absolute forecast errors, audit firm industry specialisation, firm's operating risk, and analyst quality proxies and control variables. Panel A shows the correlation matrix for the variables used in tests of H1b and H2. Absolute forecast errors (*ABSFE*) are correlated negatively and significantly with the continuous measure of auditor industry specialisation (*INDSP*)[100] and are positively and significantly associated with cash flow volatility (*CFVOL*). All

[100] In the untabulated correlation matrix, both measures of audit firm industry specialisation (continuous and dichotomous measures) are significantly related to raw or price-deflated forecast errors.

control variables are significantly associated with absolute forecast errors.[101] The correlations among control variables are generally below 0.5, with the exception that *LOSS* has a correlation with earnings per share (*EL*) of –0.5872, and absolute earnings change (*ABSECHG*) and forecast dispersion (*DISP*) has a correlation of 0.5494.

Panel B of Table 6.6 reports the correlation matrix for key variables employed in tests of H3a and H3b. As expected, absolute forecast errors (*ABSFE*) are correlated negatively and significantly with each of the analyst quality proxies. The dependent variable for tests of H3b, *DIFABSFE*, calculated based on the composite score measure is negatively correlated with *INDSP*.[102] In addition, positive and significant correlations are observed among the analyst quality proxies, with the highest correlation (0.5188) between the two experience proxies, as expected. The correlations among the analyst quality proxies are similar to those reported by Drake and Myers (2011). The variables proxied for analysts' portfolio complexity (*FFOLLOW* and *IFOLLOW*) are correlated with *ABSFE*, *INDSP* and analyst quality proxies. In the untabulated correlation matrix, there are significant correlations among my analyst quality proxies and all of the control variables. For each regression estimated, I follow prior literature (e.g. O'Brien 2007) and calculate VIFs to examine the potential impact of multicollinearity diagnosis on my models. No VIF exceeds three, suggesting no serious multicollinearity concern for my models.

---

[101] The tabulated correlation matrix includes price-deflated absolute forecast errors. In the untabulated correlation matrix, control variables are also significantly correlated with undeflated absolute forecast errors.

[102] For brevity, I only tabulate the correlation for difference in forecast errors (*DIFABSFE*) estimated based on the composite score proxy. Similar correlations are found when using other analyst quality proxies to identify the 'worst' and 'best' analysts in computing *DIFABSFE*.

Table 6.6: Correlation Matrix

Panel A Correlation Matrix for Variables Used in the Tests of H1b and H2 (BCK-Type Models)

	ABSFE	*INDSP _cont*	*CFVOL*	*DISP*	*SIZE*	*ABSECHG*	LOSS	*ZSCORE*	HORIZON	*STDROE*	*NUMEST*	*EL*
ABSFE	1											
INDSP_cont	−0.0393	1										
CFVOL	0.1863	−0.0808	1									
DISP	0.5052	0.017	0.2254	1								
SIZE	−0.3446	0.1343	−0.3006	−0.2951	1							
ABSECHG	0.7283	−0.0191	0.2468	0.5494	−0.3278	1						
LOSS	0.4538	0.0641	0.2826	0.4518	−0.3118	0.44	1					
ZSCORE	0.2341	0.1	−0.0648	0.289	−0.0029	0.2477	0.2878	1				
HORIZON	0.0538	0.0135	0.0142	0.0351	−0.0527	0.0185	0.0401	0.0186	1			
STDROE	0.2055	0.0628	0.4476	0.2754	−0.2492	0.2816	0.3944	0.1706	0.0379	1		
NUMEST	−0.0697	0.0025	−0.0253	−0.0062	0.475	−0.0474	−0.0561	−0.0144	−0.0256	−0.0208	1	
EL	−0.3347	−0.0043	−0.2964	−0.3175	0.4961	−0.3149	−0.5872	−0.1261	−0.0316	−0.3384	0.1647	1

Panel B Variables Used in the Tests of H3a and H3b

	ABSFE	*DIFABSFE _CSCORE4*	*INDSP _cont*	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE4*	*FFOLLOW*	*IFOLLOW*
ABSFE	1									
DIFABSFE _CSCORE4	0.0251	1								
INDSP _cont	−0.0236	0.002	1							
GEXP	−0.0183	0.0015	−0.0354	1						
FEXP	−0.0525	−0.0006	0.0163	0.5188	1					
BSIZE	−0.0329	0.0057	0.0485	0.0569	0.095	1				
STAR	−0.0173	0.0042	−0.0367	0.1838	0.2114	0.35	1			
CSCORE4	−0.0441	0.0011	0.0327	0.4253	0.4143	0.3876	0.1663	1		
FFOLLOW	−0.0092	0.0008	0.0917	0.2227	0.1859	−0.0846	0.1389	−0.0257	1	
IFOLLOW	−0.0107	0.0027	−0.08	0.1413	0.0531	−0.2487	−0.0626	−0.0647	0.4267	1

This table presents the correlation matrix for variables used in the tests of H1b and H2 (Panel A) and H3a and H3b (Panel B). Variable definitions are provided in Table 6.7.

Table 6.7: Variable Definitions

ABSFE	=	analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price;
ABSFE_un		analysts' absolute forecast errors (undeflated), measured as the absolute value of the difference between actual *I/B/E/S* earnings per share and forecast earnings per share;
DISP	=	forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by the beginning-of-month stock price;
DISP_eps	=	forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the absolute value of the mean EPS forecast during the period;
HORIZON (days)	=	the average number of days between mean forecast estimation date and subsequent actual earnings reporting date;
HORIZON (log)	=	the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date;
GEXP	=	the average general experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where general experience is measured as the number of years through year *t* for which an analyst *i* supplied at least one forecast for any firm;
FEXP	=	the average firm experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where firm experience is measured as number of years through year *t* for which an analyst *i* supplied at least one forecast for firm *j*;
BSIZE	=	the average brokerage size that employs analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where brokerage size is measured as number of analysts employed by a broker employing analyst *i* who follows firm *j* in year *t*;
STAR	=	the proportion of the analysts following firm *j*, during the long-horizon forecast window, who are ranked as an 'All-Star' by *II's* All-America Research Team in year *t;*
CSCORE3 (*CSCORE4*)	=	composite score, measured as the aggregate of the analysts' rankings for each individual analyst quality proxy, where the ranking is conducted within the cohort of analysts following a firm in a given year. *CSCORE3* incorporates the rankings for four proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*). CSCORE 4 incorporates the rankings for three proxies (*GEXP*, *FEXP* and *STAR*);
FFOLLOW	=	the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window;
IFOLLOW	=	the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window;
INDSP_cont	=	the continuous measure of portfolio-share industry specialisation, measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm;
INDSP_dum	=	the dichotomous measure of portfolio-share industry specialisation, equal to 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise;
SIZE (asset)	=	the natural log of total assets;
SIZE (market cap)	=	the natural log of the market value of equity;
NUMEST(number)	=	the number of analysts issuing forecasts for the firm in the 90-day window prior to reporting date;
NUMEST(log)	=	the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting;

ZSCORE	=	Zmijewski's financial distress score;
ABSACCR	=	the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets;
LOSS	=	1 if a firm reports negative earnings, 0 otherwise;
PERSIST	=	1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise;
ABSECHG	=	the absolute value of the change in annual earnings, deflated by beginning-of-month stock price;
STDROE	=	the standard deviation of return on equity over the previous five years;
EL	=	earnings per share, winsorized at 5 (-5);
CFVOL (raw)	=	cash flow volatility, measured as the 5-year standard deviation of net cash flows from operating activities deflated by average total assets;
CFVOL (log)	=	cash flow volatility, measured as the natural log of the 5-year standard deviation of net cash flows from operating activities deflated by average total assets;
DIFABSFE	=	the absolute forecast error of the 'worst' quality analyst minus the absolute forecast error of the 'best' quality analyst where the 'worst' and 'best' quality analysts are determined according to various analyst quality proxies defined above: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
ABSFE_B	=	the absolute forecast error of the 'best' quality analyst, according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*;
ABSFE_W	=	the absolute forecast error of the 'worst' quality analyst, according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*;
ANQ_B	=	the level of the quality proxy for the 'best' analyst, where the 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
DIFANQ	=	the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4;*
HORIZON_B	=	the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst, where 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*; and
DIFHORIZON	=	the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*.

6.5 Chapter Summary

In this chapter, I described the sources of data and procedures used to determine my final sample for each hypothesis test in Section 6.2. In Section 6.3, I reported the descriptive statistics for the variables employed in each model and conducted univariate comparisons of variables across clients of industry specialist audit firms and those of non-specialists for both unrestricted and propensity score matched samples. Section 6.3 contained the correlation analysis. In the following chapters, I present and analyse the results for each hypothesis test.

CHAPTER 7: RESULTS FOR TESTS OF HYPOTHESES 1A AND 1B

7.1 Introduction

The previous chapters described the measurement of key variables, regression models and the sample employed in my thesis. In this chapter, I present and analyse the results for the tests of Hypotheses 1a and 1b. Section 7.2 reports and analyses the results for the tests of Hypothesis 1a, which predicts a non-directional association between short-horizon forecast accuracy and audit firm industry specialisation. The results for the tests relating to the relationship between long-horizon forecast accuracy and audit firm industry specialisation (Hypothesis 1b) are detailed in Section 7.3. To examine the robustness of my main results, I present and discuss the sensitivity and additional test results in Sections 7.4 and 7.5. Section 7.6 concludes the chapter.

7.2 Tests of Hypothesis 1a (Short-Horizon Regressions)

Hypothesis 1a predicts that analysts' absolute short-horizon forecast errors are associated with audit firm industry specialisation. As earlier papers (BCK 2008; Payne 2008) make diametrically opposite predictions regarding the relation between auditor industry specialisation and forecast accuracy, and each of those predictions are based on plausible rationales, I do not predict a sign for *INDSP* coefficients and present two-tailed p-values adjusted for within-firm clustering using the method of Petersen (2009). Below I present and analyse the results of the regressions based on BCK-type (Section 7.2.1) and Payne-type models of short-horizon forecast errors

(Section 7.2.2). For brevity, I do not tabulate the coefficients for the industry and year indicator variables.

7.2.1 BCK-type Models

I discussed the limitations in BCK's original model in Chapter 5 and augmented the model to include an additional control variable and correction for the endogenous determination of auditor industry specialisation. The results for the original BCK model are now analysed, followed by those for the alternately specified models.

7.2.1.1 Models based on BCK

Table 7.1 reports the results for the regressions based on BCK (2008). Model BCK 1a is the original BCK's model, while Model 1b differs from the original only in that the *INDSP* is measured using a dichotomous variable.

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 HORIZON + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ZSCORE + \beta_6\, LOSS + \beta_7 ABSECHG + \beta_8 STDROE + \beta_9 EL + YEAR + \varepsilon \qquad \text{(Model BCK 1a and 1b)}$$

Models BCK 1a and 1b are well fitted ($R^2 = 0.310$).[103] While BCK report that industry specialisation is negatively associated with forecast errors for a non-Big 5 client sample, they find no significant relationship between a continuous measure of *INDSP* and forecast errors when their sample is restricted to Big N clients (β = –0.0164, two-tailed $p = 0.4715$). My re-estimation of BCK's original model over the 1989–2010 period also generates an insignificant but considerably smaller negative

[103] These R^2 statistics are considerably higher than those reported by BCK. I consider this largely arises because of the distribution of the *ABSECHG* variable. BCK do not clearly define the measurement of this variable. I used the priced-deflated change in actual EPS as reported by *I/B/E/S* (as per Payne 2008). From the descriptive statistics in BCK, it appears that they may have used the *COMPUSTAT*-defined earnings per share, shocks to which will have a noisier association with forecast errors, because analyst-defined earnings frequently 'strip out' the impact of one-off charges affecting GAAP earnings.

coefficient for *INDSP* (Model BCK 1a $\beta = -0.0013$, $p = 0.608$) in a Big N client sample. To investigate whether this difference in significance is driven by variation in the sample period, I re-estimate Model BCK 1a using BCK's sample period and find that the coefficient for *INDSP* changes little. However, if I replace the absolute change in *I/B/E/S* earnings per share (*ABSECHG*) with the absolute change in GAAP EPS from continuing operations, the coefficient for *INDSP* is –0.017 (almost identical to BCK's coefficient) and there is a significant deterioration in model fit ($R^2 = 0.254$). The coefficient for the dichotomous measure of *INDSP* (Model BCK 1b) is also negative but insignificant. All control variables are significant with the exception of the earnings level (*EL*). *STDROE* has the opposite sign to that predicted (i.e. absolute forecast errors are lower for firms with greater variability in earnings), which may reflect the mechanical correlation between this variable and *ABSECHG*. If *ABSECHG* is omitted from the regression, *STDROE* becomes positive and significant. In summary, the re-estimation of BCK's models confirms BCK's results that auditor industry specialisation is not associated with short-horizon analysts' absolute forecast errors.

7.2.1.2 BCK model with control for dispersion

In Models BCK 2a and 2b, I add price-deflated dispersion (*DISP*) as an additional control variable, as presented below.

$$ABSFE = \beta_0 + \beta_1 INDSP + \boldsymbol{\beta_2 DISP} + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + \varepsilon \qquad \text{(Model BCK 2a and 2b)}$$

In Columns BCK 2a and 2b of Table 7.1, while model fit improves significantly (increases in R^2 of 7.4 per cent), the coefficient for auditor industry specialisation is not significant in either regression. Coefficients for the other control variables are similar to those reported in Models BCK 1a and 1b.

7.2.1.3 BCK models with endogeneity correction

To control for the potentially endogenous selection of auditors, I employ a 2SLS regression for the model using the continuous measure of *INDSP* (Model 3a) and a propensity score matched sample regression for that using the dichotomous measure of *INDSP* (Model 3b). I now discuss the results of the 2SLS regressions, followed by an analysis of the propensity score matched sample results.

Under 2SLS, the first-stage regression is estimated by OLS (Model BCK 3a-1), where I obtain the predicted value for auditor industry specialisation, which replaces the actual value of industry specialisation in the second-stage model (Model BCK 3a-2). The first-stage model of the 2SLS regressions (Model Payne 3a-1) includes all of the controls from the second-stage model and two instrumental variables: the relative size of an industry (*INDRELSIZE*) and client's operating cycle (*CYCLE*). The models are as follows:

First-stage model (2SLS)

$$INDSP = \beta_0 + \beta_1 INDRELSIZE + \beta_2 CYCLE + \beta_3 DISP + \beta_4 HORIZON + \beta_5 SIZE + \beta_6 NUMEST + \beta_7 ZSCORE + \beta_8 LOSS + \beta_9 ABSECHG + \beta_{10} STDROE + \beta_{11} EL + YEAR + \varepsilon \qquad \text{(Model BCK 3a-1)}$$

Second-stage model (2SLS)

$$ABSFEadj = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + \varepsilon \qquad \text{(Model BCK 3a-2)}$$

I report the results of the second-stage of the endogeneity-corrected regressions (Model BCK 3a-2) in Column BCK 3a of Table 7.1. The results of the first-stage regression are tabulated in Appendix E. Several specification tests are conducted to examine the relevance and exogeneity of the candidate instruments. First, I examine the LM statistics, Wald-type F-statistics and partial R^2 to assess the relevance (particularly, weak identification) of my instruments. I report a large LM statistic (7228.503, $p < 0.001$), which indicates that the correlation between my candidate instruments and the endogenous regressor (*INDSP*) are statistically different from zero (i.e. not under-identified). The reported F-statistic (F = 3538.255), which is much larger than the suggested critical F-value (11.59) for cases where the number of instruments is two (Stock et al. 2002), indicates that my instruments are not weakly identified. The partial R^2 of my instruments in the first-stage regression is 38.16 per cent. Therefore, my instruments are clearly relevant to the prediction of the endogenous variable (*INDSP_cont*). I then conduct Sargan over-identification tests to examine the exogeneity of the instruments. The null hypothesis under the Sargan test is that all candidate instruments are exogenous, and the reported Sargan over-identification statistics present no evidence that the instruments for auditor specialisation are endogenous (p = 0.9511). Having established an appropriately specified model, I then estimate Durban-Wu-Hausman (DWH) tests to assess endogeneity. The DWH test of Model BCK 3a-2 generates evidence of endogeneity (p = 0.0365), indicating that OLS estimates are inconsistent. After correcting for this endogeneity, the continuous measure of auditor specialisation (*INDSP_cont*) has a negative but insignificant coefficient (β = –0.0064, p = 0.274).

To control for the impact of endogenous selection of auditors in Model BCK3b-2, which measures specialisation using a dichotomous variable, a PSM approach is employed. The base specifications for the first and second-stage models, which use the controls from the single-stage regression model as predictors of auditor industry specialisation, are presented below in Models BCK 3b-1 and 3b-2.

First-stage model (PSM)

$$INDSP = \beta_0 + \beta_1 DISP + \beta_2 HORIZON + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ZSCORE + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 STDROE + \beta_9 EL + YEAR + INDUSTRY + \varepsilon \quad \text{(Model BCK 3b-1)}$$

Second-stage model (PSM)

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 HORIZON + \beta_4 SIZE + \beta_5 NUMEST + \beta_6 ZSCORE + \beta_7 LOSS + \beta_8 ABSECHG + \beta_9 STDROE + \beta_{10} EL + YEAR + INDUSTRY + \varepsilon \quad \text{(Model BCK 3b-2)}$$

Using the first-stage regression (Model BCK 3b-1, tabulated in Appendix D) to calculate propensity scores, and imposing a one-to-one matching subject to a maximum absolute difference in propensity scores (i.e. a caliper distance) of 2.5 per cent, [104] the matched sample consists of 4,606 firm-year observations, which represents approximately 16 per cent of the unrestricted sample. As reported in Chapter 6, this sample is well balanced, with no significant differences in the means of any control variables across the treatment and control firms. I tabulate the results for tests of the matched sample in Column BCK 3b of Table 7.1, in which the coefficient for my test variable, *INDSP_dum*, is positive and insignificant (β = 0.0008, $p = 0.153$). All control variables are similar to those reported above.

---

[104] Rosenbaum and Rubin (1983) suggest that caliper size should be less than or equal to a quarter of a standard deviation of the estimated propensity score of the sample. Applying this constraint to my sample, I estimated the first-stage regression using various caliper sizes equal to or less than 0.095 (25 per cent of the standard deviation of the estimated propensity score of my sample). A caliper size of 0.025 is used as a matching criterion because this provides the largest balanced sample.

In Chapter 5, I noted that a major difference between BCK's regressions and those of Payne is that BCK do not control for industry effects. Given that both forecast errors and my measures of auditor industry specialisation are likely to be correlated with industry grouping, I conduct further untabulated analysis of the BCK-type models. I find that insignificant coefficients for auditor industry specialisation proxies can be generated in all BCK-type models either by (a) adding industry dummies representing each two-digit SIC group, (b) adding industry-year dummies, (c) adjusting client forecast accuracy by the industry-year mean of this measure or (d) adjusting all regressors by their industry-year mean.

In summary, tests based on the BCK-type models generate no evidence that auditor industry specialisation is associated with analysts' absolute short-horizon forecast errors. Thus, H1a is not supported.

Table 7.1: Short-Horizon Forecast Errors against Current Year Audit Firm Industry Specialisation (Tests of H1a using BCK-type Models)

Dependent Variable: *ABSFE*							
Models/Columns		BCK 1a	BCK 2a	BCK 3a	BCK 1b	BCK 2b	BCK 3b
		INDSP_cont (continuous measure)			*INDSP_dum (dummy measure)*		
	Pred.	BCK's Model	Price-deflated *DISP*	2SLS	BCK's Model	Price-deflated *DISP*	PSM
INDSP	?	−0.0013	0.0011	−0.0064	−0.0003	0.0002	0.0008
		(0.608)	(0.603)	(0.274)	(0.417)	(0.471)	(0.153)
DISP	+		0.6151***	0.5604***		0.6152***	0.6539***
			(<0.001)	(<0.001)		(<0.001)	(<0.001)
HORIZON	+	0.0007***	0.0005***	0.0005**	0.0007***	0.0005***	0.0018***
		(0.001)	(0.005)	(0.012)	(0.001)	(0.005)	(<0.001)
SIZE	?	−0.0011***	−0.0005***	−0.0004***	−0.0011***	−0.0005***	−0.0007***
		(<0.001)	(<0.001)	(0.001)	(<0.001)	(<0.001)	(0.003)
NUMEST	−	−0.0017***	−0.0013***	−0.0017***	−0.0017***	−0.0013***	−0.0011**
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.016)
ZSCORE	+	0.0013***	0.0003***	0.0004***	0.0013***	0.0004***	0.0005*
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.054)
LOSS	+	0.0137***	0.0069***	0.0064***	0.0137***	0.0069***	0.0061***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.1707***	0.1119***	0.1040***	0.1706***	0.1119***	0.1297***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	−0.0138***	−0.0114***	−0.0119***	−0.0137***	−0.0115***	−0.0069
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.212)
EL	+	0.0002	0.0001	0.0003*	0.0002	0.0001	−0.0003
		(0.111)	(0.416)	(0.072)	(0.119)	(0.398)	(0.446)
YEAR		yes	yes	yes	yes	yes	yes
CONSTANT		0.0189***	0.0095***	−0.0037***	0.0189***	0.0095***	0.0096***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.008)
N		35,455	29,401	28,701	35,455	29,401	4,606
R^2		0.310	0.384	0.315	0.310	0.384	0.406

Durbin-Wu-Hausman test (*p*-value)	0.0365
Sargan Over-identification test (*p*-value)	0.9511
Partial-R^2 of instruments in first-stage	0.3816
Wald F-statistics	3538.255
LM	7228.503

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H1a based on BCK's original model (Columns BCK 1a and BCK 1b), BCK's model with control for forecast dispersion (Columns BCK 2a and BCK 2b), and BCK's model with endogeneity correction using a 2SLS regression (Column BCK 3a) or a propensity score matched sample regression (Column BCK 3b).

Variable Definitions: *ABSFE* is analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price; *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation (all models suffixed 'a'), measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), equals 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise (*INDSP_dum*); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010.

7.2.2 Payne-type Models

The original Payne's (2008) models and adaptation of these models employing alternate deflators and correction for endogenous determination of industry specialist auditors were described in Chapter 5. In this section, I analyse the results for the original Payne's models first, followed by those for the alternately specified models.

7.2.2.1 Original Payne model

Table 7.2 reports the results of the Payne-type models. Models Payne 1a and 1b presented below are specified as per Payne (2008), in which audit firm industry specialisation is measured using continuous or dichotomous variables, respectively, and analyst forecast accuracy and dispersion are measured in undeflated cents per share.

$$ABSFE\_un = \beta_0 + \beta_1 INDSP + \beta_2 DISP\_eps + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ABSACCR + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 PERSIST + YEAR + INDUSTRY + \varepsilon \qquad \text{(Models Payne 1a and 1b)}$$

Columns Payne 1a and 1b of Table 7.2 present the results for the original Payne's models. The models are reasonably well fitted ($R^2 = 0.181$ in each). Consistent with Payne's results, both measures of *INDSP* are positive ($\beta = 0.0718$ for *INDSP_cont;* $\beta = 0.009$ for *INDSP_dum*), however only the dichotomous measure is significant in a two-tailed test ($p = 0.012$). The coefficient for *INDSP_cont* is smaller than that reported by Payne ($\beta = 0.158$). This possibly reflects Payne's use of the *I/B/E/S* Summary History file, rather than the Detail History file, and the resulting inclusion of stale forecasts that subtend larger forecast errors. If I follow Payne and apply unadjusted OLS standard errors (rather than firm-clustered standard errors) the coefficient for *INDSP_cont* is significant (two-tailed $p = 0.051$). All control variables are in the predicted directions and significant. *SIZE*, for which I do not predict a direction, is significantly positively associated with analysts' forecast errors, likely reflecting the higher level of undeflated EPS for large firms. Overall, the re-estimation of Payne's models provides tentative evidence of a positive relation between auditor industry specialisation and absolute undeflated forecast errors, consistent with the proposition that specialist auditors successfully constraining earnings management intended to 'meet or just beat' consensus forecasts.

7.2.2.2 Payne's model with alternative deflators

In Models Payne 2a and 2b (presented below), Payne's original model is re-estimated after deflating both the absolute forecast errors and forecast dispersion by the beginning-of-month stock price (consistent with BCK and much of the literature).[105] As undeflated forecast accuracy is mechanically affected by the level of earnings per share, which in turn is positively associated with the audit quality proxies, there is potential for spurious correlation.

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 DISP + \beta_3 SIZE + \beta_4 NUMEST + \beta_5 ABSACCR + \beta_6 LOSS + \beta_7 ABSECHG + \beta_8 PERSIST + YEAR + INDUSTRY + \varepsilon \quad \text{(Models Payne 2a and 2b)}$$

The adjusted models show that the R^2 statistics (0.383; 0.380) are more than double those of the Models Payne 1a and 1b. Saliently, the coefficient for the dichotomous measure *INDSP_dum* (Model Payne 2b) is now *negative* and significant (β = −0.0013, p = 0.009), directly contradicting the results produced by Model Payne 1b. The coefficient for the continuous measure of *INDSP* (Model Payne 2a) is also

[105] Results are qualitatively similar to those reported if I scaled forecast errors by the absolute value of actual earnings.

negative, but insignificant. Most control variables remain significant, although *SIZE* bears no relation to absolute forecast errors in these models.

7.2.2.3 Endogeneity correction in Payne's model

As above, I use 2SLS and propensity score matched sample regressions to correct for the possible endogenous selection of industry specialist audit firms, the results of which are discussed below. Column 3a of Table 7.2 reports the results of the endogeneity-corrected 2SLS regressions. While the 2SLS is employed, my instruments (*INDRELSIZE* and *CYCLE*) are both relevant (LM statistics: 1106.083, $p < 0.001$; F-statistics: 548.529), and exogenous (p-values from Sargan tests of 0.1744). There is strong evidence of endogeneity in Model Payne 3a (DWH test: $p < 0.01$), and the endogeneity-corrected coefficient for *INDSP* is now significantly negative ($\beta = -0.0148$, $p = 0.0124$).[106] Column 3b of Panel B reports the results of tests using the propensity score matched sample. The means of all covariates are insignificantly different across treatment and control firms in the matched sample and *INDSP* has an insignificant and positive coefficient ($\beta = 0.0006$, $p = 0.231$).

[106] Given the sensitivity of two-stage models to instrument choice and model specification, I conducted a series of untabulated sensitivity tests, including the substitution of clients' capital intensity for (alternately) relative industry size and operating cycle. The coefficient for *INDSP* remained negative and insignificant, but the Sargan tests rejected the null that all instruments were exogenous.

Table 7.2: Short-Horizon Forecast Errors against Current Year Audit Firm Industry Specialisation (Tests of H1a using Payne-type Models)

Dependent Variable: *ABSFE*							
Models/Columns		Payne 1a	Payne 2a	Payne 3a	Payne 1b	Payne 2b	Payne 3b
		INDSP_cont (continuous measure)			*INDSP_dum (dummy measure)*		
	Pred.	Payne's model	Price-deflated *ABSFE & DISP*	2SLS	Payne's model	Price-deflated *ABSFE & DISP*	PSM
INDSP	?	0.0718	−0.0028	−0.0148**	0.009**	−0.0013***	0.0006
		(0.147)	(0.461)	(0.0124)	(0.012)	(0.009)	(0.231)
DISP	+	0.0755***	0.621***	0.556***	0.0755***	0.622***	0.7045***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
SIZE	?	0.0175***	0.0001	0.0004***	0.0175***	0.0001	0.0002
		(<0.001)	−0.843	−0.001	(<0.001)	−0.952	(0.533)
NUMEST	−	−0.0033***	−0.0002***	−0.0002***	−0.0034***	−0.0002***	−0.0002***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSACCR	+	0.0812***	0.0024	−0.0018	0.0811***	0.0023	0.0016
		(<0.001)	(0.173)	(0.365)	(<0.001)	(0.188)	(0.682)
LOSS	+	0.0576***	0.0061***	0.0057***	0.0573***	0.0056***	0.0069***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.5440***	0.0947***	0.1080***	0.5450***	0.0945***	0.1212***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
PERSIST	−	−0.0072***	−0.0005***	−0.0006***	−0.0073***	−0.0005***	0.0007
		(<0.001)	(0.001)	(0.003)	(<0.001)	(0.001)	(0.157)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	no	yes	yes	yes
CONSTANT		0.0165	0.0018	−0.0136***	0.0176	0.0063***	0.1629**
		(0.725)	(0.496)	(<0.001)	(0.708)	(<0.001)	(0.019)
N		31,806	31,806	31,358	31,806	31,806	5,316
R^2		0.181	0.383	0.309	0.181	0.380	0.410

Durbin-Wu-Hausman test (*p*-value)	0.0008
Sargan Over-identification test (*p*-value)	0.1744
Partial-R^2 of instruments in first-stage	0.3894
Wald F-statistics	548.529
LM statistics	1106.083

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H1a based on Payne's original model (Columns Payne 1a and Payne 1b), Payne's original model with alternative deflators for forecast errors and forecast dispersion (Columns Payne 2a and 2b) and Payne's original model with endogeneity correction using a 2SLS regression (Column Payne 3a) or a propensity score matched sample regression (Column Payne 3b).

Variable Definitions: *ABSFE* is analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price (undeflated in Models Payne 1a and 1b); *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation, (all models suffixed 'a'), measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), equals 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise (*INDSP_dum*); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price (deflated by the absolute value of the mean EPS forecast during the period in Models Payne 1a and 1b); *SIZE* is the natural log of total assets; *NUMEST* is the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ABSACCR* is the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price (deflated by the natural log of total assets in Models Payne 1a and 1b); *PERSIST* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise; *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

In summary, regressions using undeflated forecast errors and undeflated dispersion generate either positive or insignificant relationships between absolute forecast errors and auditor specialisation. However, when forecast errors and dispersion are deflated by stock price, there is no evidence supporting the existence of a positive association between auditor industry specialisation and absolute forecast errors. In fact, the models using continuous measures of specialisation document a significant *negative* association between these variables.[107] This negative relationship remains when the 2SLS is used to correct for endogeneity, but an insignificant association between the dichotomous measure of industry specialisation and forecast errors is reported in the propensity score matched sample regressions.

In Table 7.3, I summarised the coefficients and p-values obtained from the BCK-type models and Payne-type models for the test variable (*INDSP*). Taken together, the results of the BCK- and Payne-type models show that the directional association between *INDSP* and short-horizon forecast accuracy is highly sensitive to model specification. Specifically, there is some evidence of a positive relation between auditor industry specialisation and absolute forecast errors when forecast errors are unadjusted, but a significantly negative relationship between these variables when forecast errors are deflated by stock price. I also note that the significance of a number of the results is sensitive to whether auditor specialisation is measured continuously or dichotomously and the vector of control variables included. These

---

[107] This juxtaposition of results plausibly reflects an imperfectly controlled relation between the level of firms' unadjusted EPS and firm size, which in turn is correlated with auditor industry specialisation. To test this contention, I added the raw level of EPS as an additional regressor to Models Payne 1a and 1b. While this additional control was significant, the coefficients for each measure of *INDSP* were not substantively affected. The coefficients (and significance levels) for *INDSP* were also similar when either the upper or lower quartile of firms ranked by size or absolute EPS were excluded from the sample, or if I exclude the upper 5 per cent of the distribution of undeflated forecast errors. I conducted similar sub-sample analysis on Models Payne 2a and 2b, and again my tabulated results were substantively repeated.

erratic results may well reflect the conflicting impacts of audit quality on short-horizon forecast errors. Conversely, higher-quality auditors produce earnings reports that are more useful for predicting future earnings, but at the same time these high-quality auditors are more effective in constraining client attempts to manage earnings in the direction of consensus forecasts. A reduction in benchmark-beating behaviour is arguably more likely to be detected using raw undeflated forecast errors because these are the measures visible to the market. While some regressions provide support for H1a, the sensitivity of these results to model specification provide little basis for confidence that short-horizon forecast accuracy is affected in either direction by auditor industry specialisation.

Table 7.3: Summary of the Results for the Test Variable from the BCK-type Models and Payne-type Models

BCK-type Models	BCK1a	BCK2a	BCK3a	BCK1b	BCK2b	BCK3b
	INDSP_cont (continuous measure)			*INDSP_dum (dummy measure)*		
INDSP	−0.0013	0.0011	−0.0064	−0.0003	0.0002	0.0008
	(0.608)	(0.603)	(0.274)	(0.417)	(0.471)	(0.153)
Payne-type Models	Payne 1a	Payne 2a	Payne 3a	Payne 1b	Payne 2b	Payne 3b
	INDSP_cont (continuous measure)			*INDSP_dum (dummy measure)*		
INDSP	0.0718	−0.0028	−0.0148**	0.009**	−0.0013***	0.0006
	(0.147)	(0.461)	(0.0124)	(0.012)	(0.009)	(0.231)

Robust p-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table summarises the coefficients and p-values for the test variable (*INDSP*) reported in Tables 7.1 and 7.2.

Variable Definitions: *INDSP* is the continuous measure (dichotomous measure) of portfolio-share audit firm industry specialisation, measured as per Equation (8) (Equation [9]).

7.3 Tests of Hypothesis 1b (Long-Horizon Regressions)

Hypothesis 1b predicts that analysts' absolute long-horizon forecast errors are negatively related to audit firm industry specialisation. Unlike H1a, this is a directional prediction because long-horizon forecasts do not obviously induce incentives for benchmark-beating behaviour. The specifications of the models used to test H1b are similar to those used to test H1a, and differ only in respect of the timing of the measurement of key variables. A negative and significant coefficient for auditor industry specialisation would be consistent with H1b. Table 7.4 presents the results for the regressions of long-horizon forecast accuracy against audit firm industry specialisation and control variables. To maintain comparability with the short-horizon results, the *p*-values reported in Table 7.4 are two-tailed, but could credibly be halved, because I have a clear directional prediction for their sign.

In Table 7.4 Panel A, I present the results of the long-horizon forecast accuracy regressions using models based on BCK. Again, all models are well fitted with R^2 statistics, ranging between 27.9 and 59.4 per cent. In all OLS models (Models BCK 1a, 1b, 2a and 2b), the coefficients for *INDSP* are negative and highly significant (two-tailed $p < 0.001$), consistent with auditor industry specialisation improving the accuracy of analysts' long-horizon forecasts. In the 2SLS regression (Column BCK 3a), the F-statistic and LM statistic are 169.263 and 872.231, respectively; while the partial R^2 statistics are approximately 15 per cent, suggesting that my instruments are relevant. The Sargan tests find no evidence that my instruments are endogenous to the error term in the second-stage regressions. The DWH test of the difference in coefficients between the efficient (OLS) and consistent (two-stage) models finds evidence of endogeneity for the continuous measure of audit specialisation (p = 0.0193); thus, the OLS estimate is inconsistent. The coefficients for *INDSP* remain negative and significant ($\beta = -0.0549$, two-tailed $p = 0.018$) after correcting for endogeneity.[108] In the matched sample regression (Column 3b), *INDSP* is significantly negative (β = –0.0019, two-tailed p = 0.091) and all control variables, except for the earnings level (*EL*), are significant.

Table 7.4 Panel B presents the results for models based on Payne (2008). All models are reasonably well fitted, with R^2 statistics ranging from 25.6 to 57.8 per cent. In Models Payne 1a and 1b, *INDSP* is negatively associated with absolute forecast errors, but is significant only in Model Payne 1a, where *INDSP* is a continuous measure (β = –0.2016, two-tailed p = 0.065). However, if I re-estimate Model Payne 1b using one-year lagged (rather than current) controls for *ABSACCR*, *ABSECHG* and *PERSIST*, *INDSP* is significantly negative (two-tailed $p < 0.05$). All control variables are significant, with the exception of forecast dispersion. As identified earlier, one of the limitations of Payne's original specification is the inconsistency between the deflators to forecast dispersion (*DISP*) and forecast accuracy (*ABSFE*).

For models in which forecast accuracy and dispersion are each deflated by stock price (Models Payne 2a and 2b), the coefficients for auditor industry specialisation are negative and significant, regardless of the measure of specialisation employed (*INDSP_cont* β = –0.0215, two-tailed $p < 0.01$; *INDSP_dum* β = –0.0051, two-tailed $p < 0.001$). Further, the appropriately specified 2SLS and matched sample regressions (Models Payne 3a and 3b) report significant negative coefficients for *INDSP* (*INDSP_cont* β = –0.1048, two-tailed $p < 0.001$; *INDSP_dum* β = –0.0021,

[108] I re-estimated Models BCK 3a and 3b by (a) using alternative sets of instruments, including client capital intensity and industry-year adjusted operating cycle, and (b) adjusting all regressors by deducting the industry-year mean from each observation. In most of these alternate formulations, I was unable to reject the possibility that at least one of the instruments was endogenous. In regressions where the Sargan test found no evidence of instrument endogeneity, the main results held.

two-tailed $p = 0.047$).[109] The DWH test ($p = 0.0024$) confirms that the presence of endogeneity affects the OLS results reported in Model Payne 2a.[110] All control variables are significant and in the predicted directions. Auditor industry specialisation is thus found to be negatively associated with absolute forecast errors in all Payne-type models that use deflated forecast errors. The only Payne-type model that generates an insignificant coefficient is Model Payne 1b; however, the endogeneity tests show that this coefficient is not a consistent estimator.

Overall, my long-horizon regressions produce results that consistently support H1b, which predicts a negative relation between audit firm industry specialisation and analysts' absolute forecast errors. These long-horizon results are far less sensitive to model specification than are those reported for short-horizon forecasts. Further, the relation between long-horizon forecast accuracy and audit firm industry specialisation is less susceptible to noise from short-term managerial benchmark-beating behaviour. I argue that this reflects the fact that long-horizon forecasts are strongly related to the quality of earnings previously reported, as other sources of timely information are relatively scant. Thus, the earnings reports audited by a higher-quality auditor are more useful for predicting future earnings, decreasing analysts' long-horizon absolute forecast errors. To test the robustness of my results further, I conduct a range of additional tests, which I describe in Sections 7.4 and 7.5.

---

[109] In BCK- or Payne-type 2SLS regressions, I lagged the control variables, *DISP*, *ABSECHG*, *ZSCORE*, *ABSACCR* and *PERSIST*, that were likely to be causally affected by audit quality. The lagged structure of the models report Sargan statistics satisfying the instrument exogeneity tests.

[110] Once again, I tested the sensitivity of my 2SLS regressions to alternate instruments, substituting a range of candidate variables (capital intensity, research and development expense and industry-adjusted operating cycle) for my main instrumental variables. The coefficients for *INDSP* remained negative and significant in every validly specified alternate model.

Table 7.4: Short-Horizon Forecast Errors against Prior Year Audit Firm Industry Specialisation (Tests of H1b)

Panel A BCK's Model

Dependent Variable: *ABSFE*							
Models/Columns		BCK1a	BCK2a	BCK3a	BCK1b	BCK2b	BCK3b
		Lag INDSP_cont (continuous measure)			*Lag INDSP_dum (dummy measure)*		
	Pred.	BCK's Model	Price-deflated *DISP*	2SLS	BCK's Model	Price-deflated *DISP*	PSM
INDSP	–	−0.0320***	−0.0295***	−0.0549**	−0.0037***	−0.0029***	−0.0019*
		(<0.001)	(<0.001)	(0.018)	(<0.001)	(<0.001)	(0.091)
DISP	+		0.3431***	0.1714**		0.3422***	0.1803
			(<0.001)	(0.016)		(<0.001)	(0.170)
HORIZON	+	0.0199***	0.0208***	0.0256***	0.0198***	0.0208***	−0.0004
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.488)
SIZE	?	−0.0037***	−0.0028***	−0.0071***	−0.0036***	−0.0028***	0.5176***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
NUMEST	–	0.0015***	0.0022***	0.0041***	0.0016***	0.0023***	0.0322***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ZSCORE	+	0.0023***	0.0012***	0.0019***	0.0022***	0.0011***	0.0019***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.003)
LOSS	+	0.0319***	0.0223***	0.0614***	0.0321***	0.0225***	0.0201***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.003)
ABSECHG	+	0.5463***	0.5160***	0.0599***	0.5459***	0.5156***	−0.0259**
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.013)
STDROE	+	−0.0396***	−0.0338***	0.0110	−0.0377***	−0.0325***	0.0011
		(<0.001)	(<0.001)	(0.177)	(<0.001)	(<0.001)	(0.204)
EL	+	−0.0001	−0.0001	−0.0005	−0.0003	−0.0002	−0.0017***
		(0.642)	(0.818)	(0.350)	(0.348)	(0.542)	(0.005)
YEAR		yes	yes	yes	yes	yes	yes
CONSTANT		−0.0672***	−0.0851***	−0.1179***	−0.0679***	−0.0855***	−0.0830**
		(0.001)	(<0.001)	(0.001)	(<0.001)	(<0.001)	(0.013)
N		32,220	25,489	21,119	32,220	25,489	4,840
R^2		0.554	0.581	0.279	0.554	0.581	0.594

Durbin-Wu-Hausman test (*p*-value)	0.0193
Sargan Over-identification test (*p*-value)	0.7378
Partial-R^2 of instruments in first-stage	0.1506
Wald F-statistics	169.263
LM statistics	872.231

Robust *p*-values of the coefficients are two-tailed reported parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H1b based on BCK's original model (Columns BCK 1a and BCK 1b), BCK's model with control for forecast dispersion (Columns BCK 2a and BCK 2b), and BCK's model with endogeneity correction using a 2SLS regression (Column BCK 3a) or a propensity score matched sample regression (Column BCK 3b).

Variable Definitions: *ABSFE* is analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price; *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation, (all models suffixed 'a'), measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), equals 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise (*INDSP_dum*); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010.

Panel B Payne's Model

Dependent Variable: *ABSFE*							
Models/Columns		Payne 1a	Payne 2a	Payne 3a	Payne 1b	Payne 2b	Payne 3b
		Lag INDSP_cont (continuous measure)			*Lag INDSP_dum (dummy measure)*		
	Pred.	Payne's model	Price–deflated *ABSFE & DISP*	2SLS	Payne's model	Price–deflated *ABSFE & DISP*	PSM
INDSP	–	–0.2016*	–0.0215***	–0.1048***	–0.0122	–0.0051***	–0.0021**
		(0.065)	(0.002)	(<0.001)	(0.291)	(<0.001)	(0.047)
DISP	+	0.0054	0.3842***	0.2290***	0.0054	0.2454***	0.2521*
		–0.591	(<0.001)	(0.002)	–0.591	(<0.001)	(0.058)
SIZE	?	0.0573***	0.0013***	0.0016***	0.0572***	0.0011***	0.0010***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.007)
NUMEST	–	–0.0056***	–0.0005***	–0.0009***	–0.0056***	–0.0005***	–0.0004***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSACCR	+	0.2835***	0.0307***	0.0060	0.2838***	0.0199***	0.0458***
		(<0.001)	(<0.001)	(0.410)	(<0.001)	(<0.001)	(<0.001)
LOSS	+	0.1829***	0.0234***	0.0744***	0.1829***	0.0206***	0.0277***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	2.7848***	0.4853***	0.0708***	2.7842***	0.4520***	0.5006***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
PERSIST	–	–0.0346***	–0.0025***	–0.0029***	–0.0346***	–0.0029***	–0.0017
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(0.107)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	no	yes	yes	yes
CONSTANT		–0.0093	0.0032	–0.0218***	–0.0103	0.0031	0.0062
		(0.962)	(0.579)	(<0.001)	(0.957)	(0.589)	(0.190)
N		27,284	27,284	21,232	27,284	27,284	5,298
R^2		0.264	0.560	0.256	0.264	0.560	0.578

Durbin-Wu-Hausman test (p-value)	0.0024
Sargan Over-identification test (p-value)	0.5845
Partial-R^2 of instruments in first-stage	0.1783
Wald F-statistics	174.513
LM statistics	940.594

Robust p-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H1b based on Payne's original model (Columns Payne 1a and Payne 1b), Payne's model with alternative deflators for forecast errors and forecast dispersion (Columns Payne 2a and 2b) and Payne's model with endogeneity correction using a 2SLS regression (Column Payne 3a) or a propensity score matched sample regression (Column Payne 3b).

Variable Definitions: *ABSFE* is analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price (undeflated in Models Payne 1a and 1b); *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation, (all models suffixed 'a'), measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), equals 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise (*INDSP_dum*); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price (deflated by the absolute value of the mean EPS forecast during the period in Models Payne 1a and 1b); *SIZE* is the natural log of total assets; *NUMEST* is the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ABSACCR* is the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price (deflated by the natural log of total assets in Models Payne 1a and 1b); *PERSIST* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise; *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

7.4 Sensitivity Analyses for Tests of Hypotheses 1a and 1b

In this section, I examine the robustness of my main results to changes in the timing of the measurement of the control variables (Section 7.4.1), alternate specifications of the propensity score matched sample regressions (Section 7.4.2) and use of Heckman treatment effects regressions as an alternate control for endogeneity (Section 7.4.3).

7.4.1 Changes in Lag-Structure of Controls

The long-horizon OLS regressions tabulated above contain control variables that depend on the realisation of the earnings being forecast (*ABSACCR*, *ABSECHG*, *PERSIST*, *DISP* and *ZSCORE*), and which may be causally affected by audit quality. I thus re-estimate the main long-horizon OLS regressions using the one-year lags of these variables (tabulated in Appendix F) and re-estimate the models with these variables excluded. The main results hold at similar significance levels, but the model fit deteriorates (e.g. none of the models report a R^2 statistic exceeding 30 per cent).

7.4.2 Alternative Thresholds for the Dichotomous Industry Specialisation Measure

To examine whether the relationship between auditor industry specialisation and long-horizon analyst forecast accuracy exists with respect to alternative thresholds from which the dichotomous variable for industry specialisation is defined, I test alternative thresholds in which *INDSP_dum* equals 1 if *INDSP_cont* exceeds between one and four times the inverse of the number of industries in existence.[111] The untabulated regressions are of similar level of model fit and generate significant negative coefficients for *INDSP*. These results suggest that my main findings are not sensitive to the thresholds used to define an industry specialist auditor.

7.4.3 Alternate Specifications of the Propensity Score Matching Regressions

I further test the sensitivity of my PSM regression results to variations in the specification of the first-stage 'matching' equation. The tabulated PSM regressions are derived from first-stage matching regressions, which estimate the conditional probability of a client hiring an industry specialist, using only the second-stage control variables included as predictors. I begin my sensitivity analysis by adding the lag of the client's operating cycle, which was used as an instrument in my 2SLS regressions and has predictive power regarding the likelihood of hiring an industry specialist auditor. My main results for tests of H1a and H1b are substantively unaffected by the inclusion of this additional predictor.

As noted in Chapter 5, matching on post-treatment variables might induce bias if the matching covariates are influenced by audit quality. To test the sensitivity of my main results to post-treatment bias, I re-estimate my short- and long-horizon regressions, using (a) predictors that are theoretically independent of the current-period auditor identity and (b) first-stage regressions that include the lagged, rather than current, values of predictors potentially subject to post-treatment bias. I replace *ZSCORE* with leverage (which is not a direct function of current profit) and replace the variables capturing client firms' earnings attributes with the equivalent cash-

[111] I test these alternative thresholds for the dichotomous industry specialisation measure for the tests of H2, H3a and H3b, and my results for these hypotheses are not affected. For brevity, I do not discuss these tests again in Chapter 8.

based variables, which are less likely to reflect a response to audit quality. I replace *LOSS* with the incidence of negative cash flows, *ABSECHG* with the absolute change in clients' cash flow from operations, *STDROE* with standard deviation of cash flow from operations, *EL* with the level of cash flow from operations and *PERSIST* with cash flow persistence (in the Payne-type models). I tabulate these alternate specifications of Model BCK 3b in Table 7.5, and show that although model fit deteriorates, the coefficients for my test variables are of similar tenor to the main results ($\beta = 0.0012$, two-tailed $p = 0.068$ in the tests of H1a; $\beta = -0.0030$, two-tailed $p = 0.093$ in the tests of H1b). I next re-estimate each model using the lagged rather than current values of *ABSECHG*, *ZSCORE* and *DISP* in both stages of Model BCK 3b, and tabulate these coefficients in Columns 3 and 4 of Table 7.5. Once more, *INDSP* has a significantly negative coefficient in the tests of H1b ($\beta = -0.0044$, two-tailed $p = 0.013$); however, the coefficient for *INDSP* in the tests of H1a is insignificant. Finally, in untabulated regressions, I re-estimate Model BCK 3b using the lagged measures of these controls in the first-stage 'matching' equation only, and excluding *ABSECHG* and *DISP* from the regressions altogether. While these regressions have lower overall model fit, the coefficients for my test variables are of similar magnitude and significance.

In summary, the evidence described above suggests that my main propensity score matched sample results are robust to variation in the specification of the prediction model and are not tainted by post-treatment bias. In the next section, I use an alternate two-stage approach (Heckman treatment effects regressions) for models with a dichotomous measure of endogenous regressor.

Table 7.5: Tests of H1a and H1b using Predictors Less Likely to be Subject to Post-Treatment Bias

Dependent Variable: *ABSFE*							
Columns		(1)	(2)			(3)	(4)
	Pred.	H1a	H1b		Pred.	H1a	H1b
INDSP	?	0.0012*	−0.0030*	*INDSP*	−	−0.0007	−0.0044**
		(0.068)	(0.093)			(0.436)	(0.013)
DISP	+	0.9740***	1.7329***	*DISP_lag*	+	0.3002***	0.1306
		(<0.001)	(<0.001)			(0.005)	(0.402)
HORIZON	+	0.0008	0.0134	*HORIZON*	+	0.0026***	0.0375***
		(0.121)	(0.255)			(<0.001)	(<0.001)
SIZE	?	−0.0007**	−0.0026***	*SIZE*	?	0.0006	0.0004
		(0.017)	(0.002)			(0.201)	(0.567)
NUMEST	−	−0.0013**	0.0003	*NUMEST*	−	−0.0037***	−0.0009
		(0.011)	(0.873)			(<0.001)	(0.459)
LEVERAGE	+	0.0052***	0.0206***	*ZSCORE_lag*	+	0.0008*	0.0018*
		(0.005)	(0.001)			(0.099)	(0.071)
NEGCF	+	0.0011	0.0102	*LOSS*	+	0.0263***	0.0721***
		(0.618)	(0.108)			(<0.001)	(<0.001)
ABSCFCHG	+	0.0000	0.0009*	*ABSECHG_lag*	+	0.0120	0.0845**
		(0.413)	(0.087)			(0.453)	(0.027)
CFVOL	+	0.0274*	0.1042**	*STDROE*	+	0.0118	0.0460**
		(0.065)	(0.020)			(0.182)	(0.013)
CF	+	−0.0167***	−0.0247	*EL*	+	−0.0017**	−0.0057***
		(0.003)	(0.127)			(0.014)	(<0.001)
YEAR		yes	yes	YEAR		yes	yes
INDUSTRY		yes	yes	INDUSTRY		yes	yes
CONSTANT		0.0018	−0.0378	*CONSTANT*		0.1996***	−0.1574***
		(0.609)	(0.576)			(<0.001)	(0.006)
N		3,134	3,196	N		3,848	4,252
R^2		0.334	0.329	R^2		0.253	0.327

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the matched sample regression results for the tests of H1a and H1b based on Model BCK 3b using predictors that are less likely to be subject to post-treatment bias. Columns (1) and (2) report the results of the regressions using predictors that are theoretically independent of the current-period auditor identity. Columns (3) and (4) report the results of the regressions using the lagged (rather than current) values of predictors that are less likely to reflect a response to audit quality.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors (an inverse function of forecast accuracy), as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *DISP* (*DISP_lag*) is forecast dispersion, measured as the (lagged value of) standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity (natural log of total assets in Payne's model); *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *LEVERAGE* is the leverage; *NEGCF* equals 1 if a firm reports negative cash flow from operations, 0 otherwise; *ABSCFCHG* is the absolute value of the change in cash flow from operations, deflated by beginning-of-month stock price; *CFVOL* is the standard deviation of cash flow from operations over the previous five years; *CF* is cash flow per share; *ZSCORE_lag* is the lagged value of Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG*_lag is the lagged value of absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

7.4.4 Heckman-type Two-Stage Regressions

My main tests employ a PSM approach to control for endogenous selection in models where audit firm industry specialisation is measured using a dichotomous variable. This method was selected because it does not rely on an assumed functional form or the validity of exclusion restrictions and provides a more direct estimate of the treatment effects (Lawrence et al. 2011). However, the limitations of the matching approach include that the generalisation of the results is dependent on the balance of the matched sample and the range of common support, and the likelihood of post-treatment bias (issues addressed in Chapter 6). Here, I employ Heckman treatment effects regressions, a method commonly employed in earlier auditing literature, to further investigate the robustness of the results generated by the PSM approach. The first-stage of Heckman regressions estimates the likelihood of appointing an industry specialist auditor using probit regression, from which I estimate the IMR and include it in the second-stage model as an additional regressor. For the first-stage Heckman regressions, I use the same instruments (industry relative size *INDRELSIZE* and clients' operating cycle *CYCLE*) as those employed in the 2SLS regressions. I now discuss the results for tests of H1a and H1b.

In Table 7.6, I present the second-stage Heckman regression results for the tests of the impact of auditor industry specialisation on short-horizon (H1a) and long-horizon (H1b) forecast accuracy, each of which is based on Model BCK 3b. I find that my test variable (*INDSP*) is insignificant in tests of H1a, but is significantly negative in tests of H1b ($\beta = -0.0044$, $p < 0.01$). Standard specification tests suggest that my instruments are both exogenous and relevant. Once more, I show that the association between auditor industry specialisation and analysts' short-horizon forecast errors is sensitive to model specification. More importantly, I demonstrate the robustness of my long-horizon findings, that audit quality decreases long-horizon forecast errors, to alternate means of controlling for the endogenous choice of auditors.

Table 7.6: Heckman Treatment-Effect Regressions for a Dichotomous Measure of Audit Firm Industry Specialisation

Dependent Variable: *ABSFE*				
Columns		(1)		(2)
	Pred.	H1a	Pred.	H1b
INDSP	?	−0.0008	–	−0.0044***
		(0.288)		(0.004)
DISP	+	0.5604***	+	0.1722***
		(<0.001)		(<0.001)
HORIZON	+	0.0005**	+	0.0253***
		(0.013)		(<0.001)
SIZE	?	−0.0004***	?	−0.0072***
		(0.001)		(<0.001)
NUMEST	–	−0.0017***	–	0.0043***
		(<0.001)		(<0.001)
ZSCORE	+	0.0004***	+	0.0018***
		(<0.001)		(<0.001)
LOSS	+	0.0065***	+	0.0613***
		(<0.001)		(<0.001)
ABSECHG	+	0.1040***	+	0.0598***
		(<0.001)		(<0.001)
STDROE	+	−0.0117***	+	0.0126***
		(<0.001)		(0.008)
EL	+	0.0003*	+	−0.0006
		(0.081)		(0.176)
IMR	?	0.0001	?	0.0002
		(0.817)		(0.853)
YEAR		yes		yes
CONSTANT		−0.0096***		−0.1016***
		(<0.001)		(<0.001)
N		28,701		21,119
R^2		0.315		0.281
Durbin-Wu-Hausman test (*p*-values)		0.7817		0.9328
Sargan Over-identification test (*p*-values)		0.8573		0.8397
Partial-R^2 of instruments in first-stage		0.2920		0.2365
Wald F-statistics		3153.174		200.59
LM		6373.849		746.676

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the second-stage Heckman treatment-effect regression results for the tests of H1a (Column 1) and H1b (Column 2) based on Model BCK 3b.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors (an inverse function of forecast accuracy), as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *IMR* is the inverse Mills ratio for endogenous auditor choice; *YEAR* is the indicator variable for each year 1989–2010.

7.5 Additional Tests for Hypotheses 1a and 1b

I conduct and report additional analyses including the use of alternate industry specialisation measures for tests of H1a and H1b in Section 7.5.1, and sub-sample analysis for H1b in Section 7.5.2.

7.5.1 Alternate Industry Specialisation Measures

While my main test results use the portfolio-share measure of auditor industry specialisation, I re-estimate my regressions using a range of alternate measures, including the market-share, weighted market-share and city-level industry specialisation measures to examine whether my main results hold. Below I provide a brief description of the alternate industry specialisation measures applied. I first use a national-level market-share-type metric to capture the effect of industry specialisation on auditor expertise. Recall that the market-share industry specialisation measure reflects an audit firm's share of the total audit fee revenue generated in an industry. While the benefits of knowledge spillover and economies of scale are likely to increase with an audit firm's market share in an industry, the negative effect of market dominance on the quality of audit outcomes may be greater for audit firms with higher-level market share (Yardley et al. 1992). Some empirical studies obtain evidence that market-share industry specialist auditor is associated with superior audit outcomes (Balsam et al. 2003; Knechel et al. 2007), while others find no significant relationship between the market-share measure of industry specialisation and various proxies for financial reporting quality (Francis et al. 2012; Minutti-Meza 2013; Boone et al. 2012). For the market-share measure, I use client's total assets to proxy audit fees to maintain consistency with the portfolio-share measure and to preserve sample size.[112] The generic form of the continuous national market-share measure of industry specialisation is as follows:

*INDSP*_market_cont	=	the sum of the square root of the total assets of the clients that an audit firm services in a particular industry divided by the sum of the square root of the total assets of all clients in that industry.	(11)

Following prior studies (Balsam et al. 2003; Reichelt and Wang 2010), the dichotomous market share is coded as 1 if the audit firm has the largest proportion of the market share in an audit industry, and 0 otherwise.[113]

I then employ the national-level weighted market-share measure proposed by Neal and Riley (2004) to estimate auditor industry specialisation. The association between weighted market-share industry specialisation and better audit outcomes is not clear because the limitations of the market-share measure may also affect the weighted market-share measure; however, it remains interesting to explore whether my results are robust to this combined measure. Following Neal and Riley (2004), this measure is the product of the continuous (dichotomous) measure of auditor market share and portfolio share:

$$INDSP\_weighted\_market = INDSP\_market * INDSP\_portfolio \quad (12)$$

While all of the industry specialisation measures I employ are estimated across the national audit markets, I also use the city-level measure because industry expertise of an audit firm is argued by some to be city-specific (Ferguson et al. 2003). Prior

[112] In the untabulated regressions, I also use audit fees to estimate market-share measure at national level. My sample size is reduced by almost 50 per cent and I obtain similar results to those estimated based on clients' total assets.
[113] I also estimate the dichotomous market-share measure, identifying specialists as firms whose market share is greater than 1.2 times the inverse of the number of Big N audit firms. My sample period for studying the national-level industry specialisation measure covers 1989–2010. There was a Big 6 for the period 1989–1997, a Big 5 for the period 1998–2001 and a Big 4 onwards. My untabulated results are qualitatively the same as those reported in Table 7.7.

studies examining city-specific expertise typically estimate the market-share measure of industry specialisation and use audit fees (rather than proxies for fees) to compute the measure since both audit firm location data and audit fee data are available post 1999 (Francis and Yu 2009; Reichelt and Wang 2010; Numan and Willekens 2012). Consistent with the literature, I estimate both the continuous and dichotomous city-level market-share specialisation measures using audit fee data obtained from Audit Analytics and conduct my sensitivity analysis over a sample period of 2000–2010.[114] The cities are defined using the U.S. Census Bureau definition of a Metropolitan Statistical Area (MSA) consistent with Francis and Yu (2009) and Reichelt and Wang (2010). The continuous city-level measure is as per Equation (13):

City-level INDSP_market_cont	=	the sum of the total audit fees an audit firm receives from its clients in an industry and in a city divided by the sum of total audit fees an audit firm receives in that city.	(13)

Following Numan and Willeken (2012), the dichotomous measure of the city-level specialisation is an indicator variable equal to 1 if the incumbent audit office is the market leader in an audit industry, and 0 otherwise.[115]

Using the above-mentioned alternate industry specialisation measures, I report the results for tests of H1a and H1b based on the matched sample regressions (Model

[114] I use client's total assets to proxy audit fees in estimating the auditor portfolio share in the main tests because fee data are not available across the full sample period. In untabulated regressions, I use audit fees rather than a proxy to estimate the national-level industry specialisation measures. My main results are unaffected, with the exception of Model Payne 2a, where the continuous measure of *INDSP* is positive and insignificant. However, the DWH statistics provide evidence of endogeneity ($p = 0.0497$), and thus, the OLS estimate obtained from Model Payne 2a is inconsistent, indicating that the 2SLS estimate ($\beta = -0.0773$, $p < 0.001$) produces convincing evidence.

[115] I also use an alternative indicator variable that recognises a city-specific industry specialist if the auditor has a market share greater than 50 per cent at city level. My untabulated results are qualitatively similar to those tabulated in Table 7.7.

BCK 3b-2) in Table 7.7.[116] In tests of H1a (Columns 1 to 3), I find that the coefficients for my test variable (*INDSP*) are insignificant in all models where the alternate industry specialisation measures are used. In tests of H1b (Columns 4 to 6), the coefficients for *INDSP* are significantly negative in models where the national-level (Column 4 β = –0.0009, two-tailed p = 0.159) and city-level market-share measures (Column 6 β = –0.0014, two-tailed p = 0.143) are used, although at lower significance levels compared to my main results. When the weighted market-share measure is used (Column 5), the coefficient for my test variable remains negative but insignificant (p = 0.193). A plausible explanation for these results is that the extent to which the weighted market-share measure captures audit quality depends on the explanatory power of the single specialisation measure. While greater market share may reflect superior technology or industry experience, this may be tempered by the fact that market share may confer a degree of monopoly power, which could reduce competition and lead to negative audit outcomes. These effects may similarly influence the complementary measure, reducing the impact of this measure of industry specialisation in improving the accuracy of analysts' prediction of earnings.

[116] The results for tests using the 2SLS regressions (Model BCK 3a) to control for the continuously endogenous industry specialisation measures are generally similar to those reported in Table 7.7.

Table 7.7: Tests of H1a and H1b using Alternate Audit Firm Industry Specialisation Measures

Dependent Variable: *ABSFE*								
Columns		(1)	(2)	(3)		(4)	(5)	(6)
			Additional Tests of H1a				*Additional Tests of H1b*	
	Pred.	National-level *INDSP*_market	National-level *INDSP*_weighted_market	City-level *INDSP*_market	Pred.	National-level *INDSP*_market	National-level *INDSP*_weighted_market	City-level *INDSP*_market
INDSP	?	−0.0002	−0.0002	0.0001	−	−0.0009	−0.0013	−0.0014
		(0.408)	(0.601)	(0.855)		(0.159)	(0.386)	(0.143)
DISP	+	0.5850***	0.5970***	0.5291***	+	0.4349***	0.3711***	0.3307***
		(<0.001)	(<0.001)	(<0.001)		(<0.001)	(0.005)	(0.002)
HORIZON	+	0.0005*	0.0010***	0.0005*	+	0.0196***	0.0170***	0.0283***
		(0.055)	(0.009)	(0.072)		(<0.001)	(<0.001)	(0.001)
SIZE	?	−0.0000	−0.0001	0.0005**	?	0.0006*	−0.0007	0.0001
		(0.916)	(0.737)	(0.013)		(0.078)	(0.224)	(0.919)
NUMEST	−	−0.0018***	−0.0016***	−0.0025***	−	−0.0008	0.0009	−0.0001
		(<0.001)	(<0.001)	(<0.001)		(0.115)	(0.336)	(0.908)
ZSCORE	+	0.0002	0.0003	0.0002	+	0.0010**	0.0019***	0.0008
		(0.118)	(0.194)	(0.233)		(0.012)	(0.004)	(0.112)
LOSS	+	0.0072***	0.0049***	0.0049***	+	0.0225***	0.0301***	0.0170***
		(<0.001)	(<0.001)	(<0.001)		(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.1184***	0.1261***	0.0848***	+	0.5147***	0.4989***	0.4920***
		(<0.001)	(<0.001)	(<0.001)		(<0.001)	(<0.001)	(<0.001)
STDROE	+	−0.0122***	−0.0051	−0.0050*	+	−0.0269***	−0.0298***	−0.0269***
		(<0.001)	(0.109)	(0.078)		(<0.001)	(0.008)	(0.002)
EL	+	−0.0002	0.0000	−0.0004**	+	−0.0019***	−0.0019***	−0.0013**
		(0.151)	(0.933)	(0.033)		(<0.001)	(0.004)	(0.027)
YEAR		yes	yes	yes		yes	yes	yes
INDUSTRY		yes	yes	yes		yes	yes	yes
CONSTANT		0.0035*	0.0030	0.0029*		−0.0971***	−0.1031***	− 0.1570***
		(0.069)	(0.479)	(0.086)		(<0.001)	(0.001)	(0.001)
N		18,076	5,616	9,682		15,415	4,824	5,334
R^2		0.394	0.416	0.347		0.574	0.580	0.555

Robust p-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the propensity score matched sample regression results for the tests of H1a (Columns 1–3) and H1b (Columns 4–6) based on Model BCK 3b using alternate audit firm industry specialisation measures. Columns (1) and (4) report the results of the regressions using the national-level market-share measure of industry specialisation. Columns (2) and (5) report the results of the regressions using the national-level weighted market-share measure of industry specialisation. Columns (3) and (6) report the results of the regressions using the city-level market-share measure of industry specialisation.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors (an inverse function of forecast accuracy), as per Equation (6b); *INDSP* equals 1 if the client is audited by a national industry specialist who has the largest market share in a two-digit SIC industry (National-level *INDSP*_market), where the market share is calculated as per Equation (11), 0 otherwise; or, if the client is audited by a national industry specialist who has the largest weighted market share in a two-digit SIC industry (National-level *INDSP*_weighted_market), where the weighted market share is calculated as per Equation (12), 0 otherwise; or, if the client is audited by a city-industry specialist who has the largest market share in a two-digit SIC industry in a city defined as the U.S. Census Bureau definition of Metropolitan Statistical Area (City-level *INDSP*_market), where the city market share is calculated as per Equation (13), 0 otherwise; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

In summary, these sensitivity analyses provide evidence consistent with my main results for the tests of H1a, suggesting that the association between audit firm industry specialisation and short-horizon forecast accuracy is affected by the models (industry specialisation measures) applied. I obtain some weak evidence to support my main results for the tests of H1b, and show that forecast accuracy increases with auditor industry specialisation when specialisation is estimated using the national- and city-level market-share measures. These results are not surprising, as I conceive auditor industry specialisation as the extent to which auditors concentrate their business by industry; the alternate measures described above do not directly capture this.

7.5.2 Other Tests for Hypothesis 1b

To aid the interpretation of results for tests of H1b, I examine whether my results are robust to re-estimation of regressions on sub-samples of firms across which it may be expected that the impact of audit quality on forecast errors will vary. For brevity, these results are discussed but are not tabulated. First, I examine whether the observed association holds for firms experiencing different financial performance. A significant negative association is documented between each measure of auditor industry specialisation and forecast accuracy for both profit-making and loss-making firms, but the coefficients for the loss-making firm regressions were of greater absolute magnitude. Given that the operations of loss-making firms are likely to be relatively volatile, this juxtaposition of results is consistent with a greater return to audit quality for these firms. Second, since auditor changes may obfuscate any causal relation between audit quality and forecast accuracy, I re-estimate the models excluding cases with auditor changes in year *t* or *t-1*. This has no substantive effect on the main results.

7.6 Chapter Summary

In this chapter, I reported and discussed the results for the tests of H1a and H1b. My results for the tests of H1a suggest that the association between audit firm industry specialisation and short-horizon forecast accuracy is highly sensitive to model specification. I find some evidence that audit firm industry specialisation is positively (negatively) associated with absolute forecast errors when forecast errors are unadjusted (deflated by stock price). These inconsistent results across different specifications of models indicate the conflicting effects of audit quality on short-horizon forecast errors. However, I demonstrated that audit firm industry specialisation decreases analysts' long-horizon absolute forecast errors, and that these results are robust to variation in model specification, supporting H1b. Sensitivity and additional analyses indicate that my main results for the tests of H1b are generally robust. These results support my argument that long-horizon forecasts are strongly related to the quality of prior period earnings, and that the accuracy of long-horizon forecasts should be improved when the quality of earnings, and thus audit quality, is increased. While these test results may provide evidence indicating the existence of a causal relationship between audit quality and forecast accuracy, I continue to examine whether this relationship varies cross-sectionally with the underlying difficulty of the forecasting task to bring more convincing evidence. I present and discuss these test results in the next chapter.

CHAPTER 8: RESULTS FOR TESTS OF HYPOTHESES 2, 3A AND 3B

8.1 Introduction

In the previous chapter, I demonstrated that the association between audit firm industry specialisation and analysts' short-horizon forecast errors is sensitive to model specification. However, I presented robust evidence in support of Hypothesis 1b, which predicts that greater industry specialisation increases the usefulness of published financial reports for the prediction of future earnings and thus improves analyst long-horizon forecast accuracy. To further investigate whether any documented empirical relationship between audit firm industry specialisation and long-horizon forecast accuracy may be causal, I test additional hypotheses (Hypotheses 2, 3a and 3b) related to cross-sectional variation in the theoretical association between audit quality and forecast accuracy. In Section 8.2, I first present and analyse the results for one source of such cross-sectional variation, which predicts that the relationship between audit firm industry specialisation and forecast accuracy is conditional on the riskiness of client firm's operating risk (H2). Sensitivity and additional tests for H2 are reported and discussed in Sections 8.3 and 8.4, respectively. Section 8.5 reports and analyses the main results for the tests of H3a and H3b, which predict that the relationship between forecast accuracy and audit firm industry specialisation is conditional on the quality of analysts issuing forecasts for a firm, measured at the firm-year level (H3a) or analyst-firm-year level (H3b). I present and discuss the sensitivity and additional test results for H3a and H3b in Sections 8.6 and 8.7. Section 8.8 concludes the chapter.

8.2 Main Tests of Hypothesis 2

Hypothesis 2 predicts that the negative association between audit firm industry specialisation and analysts' long-horizon absolute forecast errors increases in strength with the level of client firm's operating risk, consistent with a greater return to audit quality where the auditor's task complexity is greater. Recall that I use cash flow volatility (*CFVOL*) to proxy operating risk and include *CFVOL* and its interaction with auditor industry specialisation (*INDSP*CFVOL*) in the empirical model (Model 4), as described below:

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 CFVOL + \beta_3 INDSP*CFVOL + \beta_4 HORIZON + \beta_5 SIZE + \beta_6 NUMEST + \beta_7 ZSCORE + \beta_8 LOSS + \beta_9 EL + YEAR + INDUSTRY + \varepsilon \qquad \textit{(Model 4)}$$

H2 is supported if the coefficient for *INDSP*CFVOL* is significantly negative. Table 8.1 reports the results for tests of H2 for both the unrestricted and propensity score matched samples. One-tailed *p*-values are reported for *INDSP, CFVOL* and *INDSP*CFVOL*. For other variables, *p*-values are two-tailed.

Overall, the models are well fitted, with R^2 statistics of approximately 27 and 31 per cent. In the unrestricted sample (Column1), the coefficient for *INDSP* is negative and significant ($\beta = -0.0242$, $p < 0.001$), consistent with the results in Table 7.2. As expected, the coefficient for *CFVOL* is significantly positive ($\beta = 0.0084$, $p < 0.001$), suggesting that firm's operating risk is associated with greater absolute forecast errors (less accurate forecasts). The coefficient for *INDSP*CFVOL* is negative and significant ($\beta = -0.0065$, $p < 0.001$), consistent with my prediction that the negative impact of audit firm industry specialisation on absolute forecast errors increases with client's operating risk. Similarly, the matched sample regression (Column 2) reports a significantly positive coefficient for *CFVOL* ($\beta = 0.0099$, $p < 0.001$) and negative coefficients, at a lower confidence level, for *INDSP* ($\beta = -0.0190$, $p = 0.028$) and *INDSP*CFVOL* ($\beta = -0.0047$, $p = 0.045$). All control variables with the exception of *SIZE* are significant. As noted earlier in Chapter 5, *SIZE* may have confounding effects on forecast accuracy, and if such effects offset each other, it is unsurprising that *SIZE* has no incremental explanatory power in the regressions.

In summary, both the unrestricted and matched sample regressions provide evidence in support of H2, which predicts that audit firm industry specialisation has a greater impact in reducing the absolute forecast errors in analysts' prediction of earnings when client firm's operating risk is higher. These results are consistent with the effect of audit quality increasing with the difficulty of audit and forecasting tasks. The empirical support for this hypothesis provides further evidence consistent with the existence of a causal relationship between auditor industry specialisation and analyst long-horizon forecast accuracy.

Table 8.1: Long-Horizon Forecast Errors against Audit Firm Industry Specialisation and Client Firm's Operating Risk (Tests of H2)

Dependent Variable: *ABSFE*			
Columns		(1)	(2)
	Pred.	Unrestricted Sample	Matched Sample
INDSP	–	–0.0242***	–0.0190**
		(<0.001)	(0.028)
CFVOL	+	0.0084***	0.0099***
		(<0.001)	(<0.001)
*INDSP*CFVOL*	–	–0.0065***	–0.0047**
		(<0.001)	(0.045)
HORIZON	+	0.0282***	0.0328***
		(<0.001)	(0.001)
SIZE	?	0.0002	–0.0005
		(0.706)	(0.529)
NUMEST	–	–0.0020***	0.0008
		(0.002)	(0.615)
ZSCORE	+	0.0053***	0.0070***
		(<0.001)	(<0.001)
LOSS	+	0.0547***	0.0568***
		(<0.001)	(<0.001)
EL	+	–0.0053***	–0.0060***
		(<0.001)	(<0.001)
YEAR		yes	yes
INDUSTRY		yes	yes
CONSTANT		–0.0754**	–0.1198**
		(0.011)	(0.039)
N		23,558	4,080
R^2		0.267	0.307

Robust *p*-values of the coefficients for *INDSP*, *CFVOL* and *INDSP*CFVOL* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H2 using an OLS regression on an unrestricted sample (Column 1) and a propensity score matched sample (Column 2).

Variable Definitions: *ABSFE* is analysts' absolute forecast errors (an inverse function of forecast accuracy), measured as the absolute value of difference between actual *I/B/E/S* earnings per share and forecast earnings per share, deflated by beginning-of-month stock price; *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *CFVOL* is cash flow volatility, measured as the natural log of the 5-year standard deviation of net cash flows from operating activities deflated by average total assets; *INDSP*CFVOL* is the interaction between *INDSP* and *CFVOL*; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

8.3 Sensitivity Tests of Hypothesis 2

In this section, I examine the sensitivity of my main results for tests of H2 to alternately specified two-stage regressions and the use of alternate proxies for firm's operating risk. I also estimate regressions on sub-samples defined by the quartiles of operating risk proxies to examine the appropriateness of the functional form assumed in the main tests of H2.

8.3.1 Alternate Specifications of the Two-stage Regressions

My main tests use a propensity score matched sample approach to control for endogenous selection of specialist auditor. Similar to the sensitivity tests conducted for H1a and H1b, I test the robustness of my propensity score matched sample results to variations in the specification of the first-stage matching equation, and to an alternative two-stage approach that is commonly used in earlier auditing literature (Heckman treatment effects regressions) for controlling for the dichotomously endogenous regressor.

My main results for the tests of H2 are substantively unaffected when I re-estimate Model 4 by (a) adding client's operating cycle to the first-stage regression, (b) replacing the current-period *ZSCORE*, which may reflect a response to audit quality, with the lagged values of this measure, or (c) replacing *ZSCORE* with leverage, which is relatively less affected by the current-period auditor identity. For brevity, these results are untabulated.

I next examine the robustness of my tabulated results to the use of Heckman treatment effects regressions. Model 4 includes an interaction between the endogenous variable, auditor industry specialisation, and a proxy for firm's operating risk (*INDSP*CFVOL*), implying that this interaction term is also potentially affected by endogeneity. Thus, in addition to the instruments included in the 2SLS regressions (industry relative size *INDRELSIZE* and client firm's operating cycle *CYCLE*), I add interactions between *INDRELSIZE, CYCLE* and the risk proxy (*INDRELSIZE*CFVOL* and *CYCLE*CFVOL*). Consequently, while the untabulated coefficient for the interaction (*INDSP*CFVOL*) is consistent with that reported in the main tests ($\beta = -0.0059$, $p < 0.001$), the Sargan tests reject the null hypothesis that all of the instruments are strictly exogenous (Sargan test: $p = 0.0037$). I tested alternate candidate instruments, including clients' capital intensity, price-earnings ratio and stock issuance, and the interactions between these alternate candidate instruments and *CFVOL*. However, I was unable to reject the possibility that at least one of the instruments was endogenous.

In summary, the above results show that my main propensity score matched sample results are not sensitive to variations in the specification of the matching model, and are not affected by post-treatment bias. For tests of H2 using the Heckman treatment effects regressions, the coefficient for the test variable is consistent with my main results; however, the inability to confirm the exogeneity of the instruments means that these findings should be treated with caution.

8.3.2 Alternate Proxies for Firm's Operating Risk

My results for the tests of H2 suggest that industry specialist auditors play a greater role in improving forecast accuracy when firm's operating risk, proxied by cash flow volatility, is higher. In this section, I re-estimate my tests of H2 using alternate proxies for firm's operating risk. I first use the untransformed cash flow volatility measure to examine whether my main results are sensitive to the transformation applied. I next use the innate accrual quality measure, developed by Francis et al. (2005a), to proxy client firm's operating risk. The innate accrual quality measure captures the extent to which firm fundamentals explain cross-sectional variation in accrual estimation errors (Francis et al. 2005a). This measure should be systematically related to firm and industry characteristics and be relatively independent of managerial choices. Poor innate accrual quality affects the mapping of accounting earnings into cash flow (Francis et al. 2005a), which is therefore likely to affect the precision of information available for the prediction of a firm's future cash flow and earnings (closely related to my definition of firm's operating risk). Therefore, greater analysts' absolute forecast errors should be associated with poorer quality of clients' innate accruals.

Using McNichols's (2002) modification of the Dechow and Dichev's (2002) accrual quality model, Francis et al. (2005a) model innate accrual quality as a function of firm size, the standard deviations of cash flow from operations, the standard deviation of sales revenues, firm's operating cycle and the incidence of negative earnings. A larger value of the fitted value indicates poorer innate accrual quality. The innate accrual quality measure is thus a more broadly based proxy for firm's underlying operating risk than is cash flow volatility. However, a potential limitation of the innate accrual quality measure is that the measure includes variables that may represent a response to audit quality, and which may thus bias against finding a significant interaction between auditor industry specialisation and risk.[117] The estimation of the innate accrual quality model is detailed in Appendix G.

Table 8.2 reports the results of the tests of H2 using these alternate risk proxies. All models are well fitted.[118] In the unrestricted samples (Columns 1 and 2), the coefficients for the risk proxies are positive and significant as expected, and those for the interaction terms are negative and significant (*CFVOL_raw:* $\beta = -0.1234$, $p < 0.001$; *INNATEAQ:* $\beta = -0.1659$, $p < 0.001$). In the matched samples (Columns 3and 4), the coefficient remains significantly negative for *CFVOL_raw*, but is negative and insignificant ($p = 0.135$) when *INNATEAQ* is examined. The results for the *INNATEAQ* measure should be interpreted with caution because the main effect (*INDSP*) is positive and significant in the unrestricted sample. While *INDSP* is in the predicted direction (negative) in the matched sample, the interaction term is insignificant. In summary, while I obtain strong evidence to reinforce my main results using the untransformed cash flow volatility measure, my sensitivity analyses based on the innate accrual quality provide less convincing support. This may reflect the fact that the *INNATEAQ* measure may itself represent a response to audit quality. For example, high-quality auditors may constrain the variation in firm's revenues, resulting in lower estimated *INNATEAQ*. The data requirements for estimating *INNATEAQ* also effectively filter younger firms, which are typically riskier.

[117] I re-computed innate accrual quality after excluding those factors that may be affected by auditor identity (i.e. the standard deviation of sales revenues and incidence of negative earnings). Untabulated regressions using this accrual quality measure generate results similar to those reported in Table 8.2.
[118] When accrual quality is used to measure firm's risk, my sample decreases to 17,358, compared to 23,558 observations in a model using cash flow volatility. The fall in sample size is largely due to the requirement for data to compute changes in working capital accounts.

Table 8.2: Tests of H2 using Alternate Proxies for Client Firm's Operating Risk

Dependent Variable: *ABSFE*

Columns		(1)	(2)	(3)	(4)
		Unrestricted Samples		Matched Samples	
	Pred.	*CFVOL_raw*	*INNATEAQ*	*CFVOL_raw*	*INNATEAQ*
INDSP	?	0.0028*	0.0057**	0.0036	−0.0017
		(0.063)	(0.018)	(0.141)	(0.364)
RISK	+	0.1406***	0.1395***	0.1915***	0.2768***
		(<0.001)	(<0.001)	(<0.001)	(0.002)
*INDSP*RISK*	−	−0.1234***	−0.1659***	−0.1466**	−0.1078
		(<0.001)	(<0.001)	(0.018)	(0.135)
HORIZON	+	0.0280***	0.0251***	0.0524***	0.0364***
		(<0.001)	(<0.001)	(<0.001)	(0.006)
SIZE	?	−0.0001	−0.0004	−0.0001	−0.0002
		(0.825)	(0.423)	(0.883)	(0.793)
NUMEST	−	−0.0017***	−0.0007	−0.0012	0.0010
		(0.005)	(0.276)	(0.378)	(0.544)
ZSCORE	+	0.0052***	0.0059***	0.0054***	0.0073***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
LOSS	+	0.0547***	0.0558***	0.0584***	0.0492***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
EL	+	−0.0053***	−0.0059***	−0.0057***	−0.0067***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
YEAR		yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes
CONSTANT		−0.1065***	−0.1323***	−0.2305***	−0.1885**
		(<0.001)	(<0.001)	(0.002)	(0.014)
N		23,558	17,358	3,900	2,804
R^2		0.267	0.273	0.304	0.299

Robust *p*-values of the coefficients for *INDSP*, *RISK* and *INDSP*RISK* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H2 using alternate proxies for client's operating risk. Columns (1) and (3) report the results of the regressions using the untransformed cash flow volatility proxy on an unrestricted sample and a propensity score matched sample. Columns (2) and (4) report the results of the regressions using the innate accrual quality proxy on an unrestricted sample and a propensity score matched sample.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors (an inverse function of forecast accuracy), as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *RISK* is cash flow volatility, measured as the standard deviation of the client's last five years' operating cash flows, deflated by total assets (*CFVOL*_raw) in Columns 1 and 3; or the innate accrual quality (*INNATEAQ*), as per Equation (13) included in Appendix G, in Columns 2 and 4 (*INNATEAQ*); *INDSP*RISK* is the interaction between *INDSP* and *RISK*; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

8.3.3 Sub-sample Tests

To further examine the robustness of the test results of H2, I estimate regressions of absolute forecast errors against auditor industry specialisation within the sub-samples defined by the quartiles of the operating risk proxy. While the interaction-based model used in my main tests (Model 4) constrains the coefficients for the control variables to be constant at all levels of risk, the sub-sample approach does not impose this constraint. The sub-sample results for the sensitivity tests of H2 are presented in Table 8.3. Client firms are grouped into quartiles based on *CFVOL*, with Quartile 1 (4) representing the sub-sample of clients with the lowest (highest) level of risk. The coefficients for *INDSP* are negative and significant across all sub-samples, among which the coefficient for *INDSP* is most negative in the 4th Quartile ($\beta = -0.0171$, $p < 0.001$).[119] Thus, while the functional form assumed in the main tests may be imperfect, this does not appear to compromise the test results.

[119] I also estimate regressions within quartiles of forecast dispersion. I find that the coefficients for auditor industry specialisation (*INDSP*) are most negative in the upper two quartiles of forecast dispersion.

Table 8.3: Sub-Sample Tests of H2

Dependent Variable: *ABSFE*					
Columns		(1)	(2)	(3)	(4)
	Pred.	Q=1 (lowest *CFVOL*)	Q=2	Q=3	Q=4 (Highest *CFVOL*)
INDSP	–	−0.0052***	−0.0064***	−0.0079***	−0.0171***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
HORIZON	+	0.0166***	0.0146**	0.0241**	0.0758***
		(0.001)	(0.011)	(0.040)	(<0.001)
SIZE	?	−0.0014***	−0.0012**	0.0008	0.0032***
		(0.002)	(0.048)	(0.298)	(0.002)
NUMEST	–	0.0001	−0.0001	−0.0020*	−0.0040**
		(0.869)	(0.934)	(0.084)	(0.015)
ZSCORE	+	0.0047***	0.0060***	0.0061***	0.0043***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
LOSS	+	0.0445***	0.0613***	0.0606***	0.0494***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
EL	+	−0.0038***	−0.0041***	−0.0048***	−0.0076***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
YEAR		yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes
CONSTANT		−0.0590**	−0.0217	−0.0972	−0.3829***
		(0.036)	(0.542)	(0.152)	(<0.001)
N		5,890	5,889	5,890	5,889
R^2		0.228	0.263	0.277	0.206

Robust p-values of the coefficients for *INDSP*, *CFVOL* and *INDSP*CFVOL* are one-tailed reported, and for others are two-tailed reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

This table presents the results for the tests of H2 within the sub-samples defined by the quartiles of the operating risk proxy, cash flow volatility. Columns 1 to 4 report the results for the sub-sample of clients with the lowest level of risk to highest level of risk.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors (an inverse function of forecast accuracy), as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *CFVOL* is cash flow volatility, measured as the 5-year standard deviation of net cash flows from operating activities deflated by average total assets; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

8.4 Additional Tests of Hypothesis 2

My main results for the tests of H2 show that auditor portfolio-share industry specialisation decreases analysts' absolute forecast errors to a greater extent when client's operating risk is higher. In Chapter 7, I described three alternate industry specialisation measures (national-level market-share, national-level weighted market-share and city-level market-share measures) and argue that these measures may also capture some aspects of auditor expertise that may lead to superior audit outcomes, such as improvements in the usefulness of financial reports for analysts' prediction of future earnings. In this section, I re-estimate tests of H2 using alternate industry specialisation measures. For brevity, I report only the matched sample results in Table 8.4. My main results are robust to models using national-level market and weighted market-share measures of auditor industry specialisation. The coefficients for *INDSP*CFVOL* are significantly negative (market share $\beta = -0.0023$, $p = 0.076$; weighted market share $\beta = -0.0058$, $p = 0.026$), while the coefficients for the main effects are in the predicted directions and are significant. However, I find no support for H2 when using the city-level market-share measure of auditor industry specialisation.

Table 8.4: Tests of H2 using Alternate Audit Firm Industry Specialisation Measures

Dependent Variable: *ABSFE*				
Columns		(1)	(2)	(3)
	Pred.	National-level *INDSP_market*	National-level *INDSP_weighted_market*	City-level *INDSP_market*
INDSP	–	–0.0089*	–0.0233**	–0.0029
		(0.063)	(0.014)	(0.352)
CFVOL	+	0.0070***	0.0087***	0.0022*
		(<0.001)	(<0.001)	(0.087)
*INDSP*CFVOL*	–	–0.0023*	–0.0058**	–0.0002
		(0.076)	(0.026)	(0.464)
HORIZON	+	0.0270***	0.0325***	0.0354***
		(<0.001)	(0.001)	(0.002)
SIZE	?	0.0006	–0.0004	0.0003
		(0.296)	(0.609)	(0.710)
NUMEST	–	–0.0024***	0.0007	–0.0014
		(0.002)	(0.640)	(0.260)
ZSCORE	+	0.0054***	0.0069***	0.0040***
		(<0.001)	(<0.001)	(<0.001)
LOSS	+	0.0537***	0.0571***	0.0450***
		(<0.001)	(<0.001)	(<0.001)
EL	+	–0.0059***	–0.0059***	–0.0041***
		(<0.001)	(<0.001)	(<0.001)
YEAR		yes	yes	yes
INDUSTRY		yes	yes	yes
CONSTANT		–0.0899**	–0.1308**	–0.1922***
		(0.011)	(0.025)	(0.004)
N		14,514	4,080	5,194
R^2		0.262	0.307	0.245

Robust *p*-values of the coefficients for *INDSP*, *CFVOL* and *INDSP*CFVOL* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the propensity score matched sample regression results for the tests of H2 based on Model (4) using alternate audit firm industry specialisation measures. Columns (1) reports the results of the regression using the national-level market-share measure of industry specialisation. Columns (2) reports the results of the regression using the national-level weighted market-share measure of industry specialisation. Columns (3) reports the results of the regression using the city-level market-share measure of industry specialisation.

Variable Definitions: *ABSFE* is the analysts' absolute earnings forecast errors, measured as per Equation (6b); *INDSP* equals 1 if the client is audited by a national industry specialist who has the largest market share in a two-digit SIC industry (National-level *INDSP*_market), where the market share is calculated as per Equation (11), 0 otherwise; or, if the client is audited by a national industry specialist who has the largest weighted market share in a two-digit SIC industry (National-level *INDSP*_weighted_market), where the weighted market share is calculated as per Equation (12), 0 otherwise; or, if the client is audited by a city-industry specialist who has the largest market share in a two-digit SIC industry in a city defined as the U.S. Census Bureau definition of Metropolitan Statistical Area (City-level *INDSP*_market), where the city market share is calculated as per Equation (13), 0 otherwise; *CFVOL* is cash flow volatility, measured as the natural log of the 5-year standard deviation of net cash flows from operating activities deflated by average total assets; *INDSP*CFVOL* is the interaction between *INDSP* and *CFVOL*; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

8.5 Main Tests of Hypotheses 3a and 3b

Another potential source of predictable variation in the effectiveness of audit quality in improving forecast accuracy derives from the quality of analysts covering a client firm. I argue that, if my earlier findings regarding the relationship between long-horizon forecast accuracy and audit firm industry specialisation actually derive from audit quality rather than a spurious correlation, auditor industry specialisation should be observed to have a greater impact on the accuracy of forecasts made by lower-quality analysts. I develop separate hypotheses regarding the relationship between auditor industry specialisation, analyst quality and forecast accuracy on a firm-year basis (Hypothesis 3a), and on an analyst-firm-year basis (Hypothesis 3b). Empirical support for these hypotheses would provide further evidence consistent with the existence of a causal relationship between audit quality and analyst forecast accuracy. Further, support for these hypotheses provides evidence regarding the extent to which audit quality contributes to the achievement of the stated objectives of financial report (i.e. usefulness of the financial reports to a greater number of financially competent users).

8.5.1 Main Tests of Hypothesis 3a

Hypothesis 3a predicts that the negative association between auditor industry specialisation and analysts' absolute long-horizon forecast errors decreases with the average quality of the analysts covering client firms. Since H3a refers to a firm-year-level prediction, I test this hypothesis using proxies for the average quality of the analysts following a firm in a given year. Model 5a includes every specific analyst quality proxy and their interactions with industry specialisation (*INDSP*). Model 5b includes (alternatively) each of the two composite measures of analyst quality and their interactions with *INDSP*. In addition to the control variables included in Models BCK 1a and 1b, I also control for analysts' portfolio complexity, proxied by the average number of firms and industries followed by an analyst (*FFOLLOW* and *IFOLLOW*). Models 5a and 5b are described below:

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 GEXP + \beta_3 INDSP{*}GEXP + \beta_4 FEXP + \beta_5 INDSP{*}FEXP + \beta_6 BSIZE + \beta_7 INDSP{*}BSIZE + \beta_8 STAR + \beta_9 INDSP{*}STAR + \beta_{10} FFOLLOW + \beta_{11} IFOLLOW + \beta_{12} DISP + \beta_{13} HORIZON + \beta_{14} SIZE + \beta_{15} NUMEST + \beta_{16} ZSCORE + \beta_{17} LOSS + \beta_{18} ABSECHG + \beta_{19} STDROE + \beta_{20} EL + INDUSTRY + \varepsilon \quad \text{(Model 5a)}$$

$$ABSFE = \beta_0 + \beta_1 INDSP + \beta_2 CSCORE1(CSCORE2) + \beta_3 INDSP{*}CSCORE1(INDSP{*}CSCORE2) + \beta_4 FFOLLOW + \beta_5 IFOLLOW + \beta_6 DISP + \beta_7 HORIZON + \beta_8 SIZE + \beta_9 NUMEST + \beta_{10} ZSCORE + \beta_{11} LOSS + \beta_{12} ABSECHG + \beta_{13} STDROE + \beta_{14} EL + INDUSTRY + \varepsilon \quad \text{(Model 5b)}$$

While the main effect of audit firm industry specialisation on absolute forecast errors (*INDSP*) is expected to be negative, H3a is supported if the interactions between the analyst quality proxies and auditor industry specialisation are positive and significant.

Table 8.5 presents the results for tests of H3a. Panel A provides the results for the unrestricted samples using OLS with standard errors adjusted for firm and year clustering, while Panel B presents the results for the propensity score matched samples. One-tailed *p*-values are reported in all of the following tables for the variables for which I have a directional prediction; otherwise, the *p*-values are two-tailed.

In Column 1 of Panels A and B, I report the results of tests in which each of the singular analyst quality proxies are interacted with auditor industry specialisation. The R^2 statistics are 55.8 per cent in the unrestricted sample (Panel A) and 53.7 per cent in the propensity score matched sample (Panel B), suggesting that the models are well fitted. In both samples, the main effect for *INDSP* is negative and significant ($p <= 0.001$), consistent with the results in Table 7.2.[120] The interaction terms for auditor industry specialisation and each analysts' expertise proxy are positive and significant in the unrestricted sample (*INDSP*GEXP* $\beta = 0.0004$, $p = 0.032$; *INDSP*FEXP* $\beta = 0.0006$, $p = 0.053$; *INDSP*BSIZE* $\beta = 0.0003$, $p < 0.01$; *INDSP*STAR* $\beta = 0.0107$, $p < 0.01$). This is consistent with my hypothesis that audit quality becomes less important to average forecast accuracy when the average quality of analysts increases. The coefficients for *INDSP*GEXP* and *INDSP*BSIZE* are also in the predicted directions and are significant in the propensity score matched samples; however, the interaction between auditor industry specialisation and analysts' firm-specific forecasting experience (*INDSP*FEXP*), and the interaction between auditor industry specialisation and the analysts' 'All-Stars' status (*INDSP*STAR*), are not significant in the matched sample. I suggest caution in interpreting any of the results of tests of H3a using the forecasting experience proxies and the mean 'All-Star' proxy, as the main effects for these variables (*GEXP*, *FEXP* and *STAR*) are insignificant, suggesting that these metrics applied in isolation may not be strong proxies for the average quality of analysts following a firm.

The coefficients for control variables imported from tests of H1b are similar to those reported in Chapter 7. Of the additional controls included in the tests of H3a, the number of industries followed by an analyst is increasing with absolute forecast errors (*IFOLLOW* $\beta = 0.0004$, $p <= 0.001$) for the unrestricted sample. This result

[120] The availability of 'All-Star' analyst data restricts the samples used to test H3a to the period 1993–2010. If this shorter sample period is used for tests of H1b, I generate similar results to those tabulated in Chapter 7 (*INDSP_dum* $\beta = -0.0018$, $p = 0.035$ in the unrestricted sample, $\beta = -0.0024$, $p = 0.026$ in the matched sample). This regression is tabulated in Appendix H.

indicates that 'busy' analysts perform worse in forecasting earnings, consistent with prior argument and evidence (Clement 1999). However, there is no association between either of the portfolio complexity measures and forecast errors in the matched sample.

In Columns 2 and 3 of Table 8.5, I present the results for the tests of H3a, in which analyst quality is proxied by alternate composite score measures. I ranked all analysts who were active within a firm's long-horizon forecast window according to each of the quality proxies and summed the rankings across the four quality proxies for each analyst (i.e. general experience, firm-specific experience, brokerage size and 'All-Star' status). *CSCORE1* is the average value of total rankings of the analysts following a client firm in a given year. *CSCORE2* is similarly defined, but brokerage size is excluded from the aggregation. The models based on the composite score are well fitted (R^2 statistics range from 53.5 to 55.9 per cent). Again, the main effect of *INDSP* on forecast errors is negative and significant in both the unrestricted and matched samples. In the tests based on the unrestricted samples (Panel A, Columns 2 and 3), the interaction terms *INDSP*CSCORE1* and *INDSP*CSCORE2* are positive but insignificant ($p = 0.114$; $p = 0.208$, respectively). However, when I estimate the regressions after matching samples of treatment and control firms with inherently similar characteristics (Panel B Columns 2 and 3), I obtain evidence of significant positive coefficients for the interactions between auditor industry specialisation and the composite analyst quality measures (p-values of 0.022).

Table 8.5: Long-Horizon Forecast Errors against Audit Firm Industry Specialisation and Analyst Quality (Tests of H3a)
Panel A OLS Regressions on Unrestricted Samples

Dependent Variable: *ABSFE*				
Columns		(1)	(2)	(3)
	Pred.	Singular Proxy	*CSCORE1*	*CSCORE2*
INDSP	–	–0.0089***	–0.0031**	–0.0026**
		(<0.001)	(0.014)	(0.035)
GEXP	–	–0.0002		
		(0.109)		
*INDSP*GEXP*	+	0.0004**		
		(0.032)		
FEXP	–	–0.0000		
		(0.421)		
*INDSP*FEXP*	+	0.0006*		
		(0.053)		
BSIZE	–	–0.0002		
		(0.119)		
*INDSP*BSIZE*	+	0.0003***		
		(0.003)		
STAR	–	–0.0000		
		(0.498)		
*INDSP*STAR*	+	0.0107***		
		(0.002)		
CSCORE	–		–0.0020***	–0.0024***
			(0.005)	(0.006)
*INDSP*CSCORE*	+		0.0005	0.0004
			(0.114)	(0.208)
FFOLLOW	?	–0.0000	–0.0000	–0.0000
		(0.587)	(0.561)	(0.567)
IFOLLOW	?	0.0004***	0.0004***	0.0005***
		(0.001)	(<0.001)	(<0.001)
DISP	+	0.3898***	0.3998***	0.3990***
		(<0.001)	(<0.001)	(<0.001)
HORIZON	+	0.0189***	0.0206***	0.0207***
		(<0.001)	(<0.001)	(<0.001)
SIZE	?	0.0001	0.0005	0.0004
		(0.808)	(0.340)	(0.401)
NUMEST	–	–0.0019***	–0.0012*	–0.0012*
		(<0.001)	(0.087)	(0.076)
ZSCORE	+	0.0010*	0.0011*	0.0011*
		(0.083)	(0.059)	(0.062)
LOSS	+	0.0186***	0.0191***	0.0191***
		(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.4928***	0.4953***	0.4953***
		(<0.001)	(<0.001)	(<0.001)
STDROE	+	–0.0216***	–0.0225***	–0.0226***
		(<0.001)	(<0.001)	(<0.001)
EL	+	–0.0020***	–0.0014***	–0.0014***
		(<0.001)	(0.001)	(<0.001)
INDUSTRY		yes	yes	yes
CONSTANT		–0.1132***	–0.0905***	–0.1245***
		(<0.001)	(<0.001)	(<0.001)
N		22,742	22,742	22,742
R^2		0.558	0.559	0.559

Robust *p*-values of the coefficients for *INDSP*, *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE*, *INDSP*GEXP*, *INDSP*FEXP*, *INDSP*SIZE*, *INDSP*STAR* and *INDSP*CSCORE* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

Panel B OLS Regressions on Propensity Score Matched Samples

Dependent Variable: *ABSFE*				
Columns		(1)	(2)	(3)
	Pred.	Singular Proxies	*CSCORE1*	*CSCORE2*
INDSP	–	–0.0164***	–0.0093***	–0.0093***
		(0.001)	(<0.001)	(0.001)
GEXP	–	–0.0005		
		(0.223)		
*INDSP*GEXP*	+	0.0011**		
		(0.042)		
FEXP	–	0.0006		
		(0.235)		
*INDSP*FEXP*	+	0.0002		
		(0.383)		
BSIZE	–	–0.0008**		
		(0.013)		
*INDSP*BSIZE*	+	0.0008**		
		(0.018)		
STAR	–	0.0086		
		(0.15)		
*INDSP*STAR*	+	0.0107		
		(0.109)		
CSCORE	–		–0.0034***	–0.0042**
			(0.008)	(0.013)
*INDSP*CSCORE*	+		0.0020**	0.0026**
			(0.022)	(0.022)
FFOLLOW	?	0.0000	0.0001	0.0001
		(0.911)	(0.307)	(0.270)
IFOLLOW	?	0.0001	0.0002	0.0002
		(0.633)	(0.493)	(0.437)
DISP	+	0.3225	0.3268	0.3256
		(0.131)	(0.133)	(0.133)
HORIZON	+	0.0107	0.0141*	0.0141
		(0.204)	(0.099)	(0.104)
SIZE	?	–0.0003	0.0004	0.0003
		(0.762)	(0.638)	(0.704)
NUMEST	–	–0.0011	–0.0002	–0.0002
		(0.367)	(0.912)	(0.888)
ZSCORE	+	0.0011	0.0013	0.0013
		(0.332)	(0.253)	(0.260)
LOSS	+	0.0213***	0.0224***	0.0224***
		(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.4780***	0.4811***	0.4807***
		(<0.001)	(<0.001)	(<0.001)
STDROE	+	–0.0193*	–0.0193*	–0.0195*
		(0.080)	(0.070)	(0.067)
EL	+	–0.0023**	–0.0015	–0.0015
		(0.021)	(0.194)	(0.194)
INDUSTRY		yes	yes	yes
CONSTANT		–0.0582	–0.0746	–0.0747
		(0.206)	(0.121)	(0.126)
N		3,686	3,686	3,686
R^2		0.537	0.535	0.535

Robust *p*-values of the coefficients for *INDSP*, *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE*, *INDSP*GEXP*, *INDSP*FEXP*, *INDSP*SIZE*, *INDSP*STAR* and *INDSP*CSCORE* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the tests of H3a using OLS regressions on unrestricted samples (Panel A) and propensity score matched samples (Panel B). Column 1 of Panel A and B reports the results of the regressions using the singular analyst quality proxies. Columns 2 and 3 report the results of the regressions using two composite measures of analyst quality.

Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *GEXP* is the average general experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where general experience is measured as the number of years through year *t* for which an analyst *i* supplied at least one forecast for any firm; *FEXP* is the average firm experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where firm experience is measured as number of years through year *t* for which an analyst *i* supplied at least one forecast for firm *j*; *BSIZE* is the average brokerage size that employs analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where brokerage size is measured as number of analysts employed by a broker employing analyst *i* who follows firm *j* in year *t*; *STAR* is the proportion of the analysts following firm *j*, during the long-horizon forecast window, who are ranked as an 'All-Star' by *II's* All-America Research Team in year *t*; *CSCORE1* (*CSCORE2*) is the composite score, measured as the average of the total ranks for analysts following a firm where the ranking is conducted according to each individual analyst quality proxy within the long-horizon forecast window; *CSCORE1* incorporates the rankings for four proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*), *CSCORE2* incorporates the rankings for three proxies (*GEXP*, *FEXP* and *STAR*); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *INDSP*GEXP* is the interaction between *INDSP* and *GEXP*; *INDSP*FEXP* is the interaction between *INDSP* and *FEXP*; *INDSP*BSIZE* is the interaction between *INDSP* and *BSIZE*; *INDSP*STAR* is the interaction between *INDSP* and *STAR*; *INDSP*CSCORE* is the interaction between *INDSP* and *CSCORE1* or *CSCORE2*; *FFOLLOW* is the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *IFOLLOW* is the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

In summary, there is mixed evidence in support of H3a. I report strong evidence in support of H3a, which argues that the relationship between audit firm industry specialisation and analysts' absolute forecast errors is stronger when the average quality of analysts (proxied by analysts' general experience or analysts' employer size) covering a firm is lower. However, when analyst quality is proxied by firm-specific experience or 'All-Star' status, H3a is only supported in the unrestricted sample. Further, results from the tests using the two composite measures of analyst quality support H3a when applied to matched samples of client firms with similar propensities to be audited by an industry specialist. This evidence is generally consistent with, but obviously not proof of, the existence of a causal relationship between auditor industry specialisation and the usefulness of published earnings for predicting future performance. In the next section, I examine this further but focus on the impact of auditor industry specialisation on the forecasting performance of *individual* analysts following a given firm.

8.5.2 Main Tests of Hypothesis 3b

Hypothesis 3b predicts that audit firm industry specialisation reduces the difference in analysts' absolute forecast errors between the 'worst' and 'best' quality analysts following a client firm in a given year. In the tests of H3b, the difference in absolute forecast errors between the 'worst' and 'best' analysts (identified separately for each singular analyst quality proxy) is regressed on auditor industry specialisation and the control variables, as per Model 6:

$$\begin{aligned} DIFABSFE = \; & \beta_0 + \beta_1 INDSP + \beta_2 ABSFE\_B + \beta_3 DIFANQ + \\ & \beta_4 ANQ\_B + \beta_5 HORIZON\_B + \beta_6 DIFHORIZON + \\ & \beta_7 DISP + \beta_8 SIZE + \beta_9 NUMEST + \beta_{10} ZSCORE + \\ & \beta_{11} LOSS + \beta_{12} ABSECHG + \beta_{13} STDROE + \beta_{14} EL + \\ & YEAR + INDUSTRY + \varepsilon \end{aligned} \qquad \text{(Model 6)}$$

H3b is supported if auditor industry specialisation (*INDSP*) is significantly negative.

Table 8.6 presents the results for the unrestricted (Panel A) and matched sample tests (Panel B) of H3b. The regressions reported in Table 8.6 differ with respect to the means by which the 'worst quality' and 'best quality' analysts are identified. The first four columns relate to differences in analysts' general forecasting experience (Column 1), firm-specific forecasting experience (Column 2), employing brokerage size (Column 3) and 'All-Star' status (Column 4), while the last two columns report regressions in which the 'best' and 'worst' analysts following a particular firm are identified using the two composite score measures of analyst quality, *CSCORE3* (Column 5) and *CSCORE4* (Column 6). As I have a directional prediction for the signs of auditor industry specialisation (*INDSP*), one-tailed *p*-values for *INDSP* are reported in all of the following tables.

The models used to test H3b are less well fitted than those used to test my other hypotheses (R^2 statistics reported in Table 8.6 range from 4.9 to 7.6 per cent). This may be because most of my control variables are defined on a firm-year basis and are likely to have a lower explanatory power regarding the differences in individual analyst forecast accuracy than they do for average level of forecast accuracy.

Columns 1 and 2 of Table 8.6 present the results for the tests of H3b, where analysts' general (*GEXP*) and firm-specific forecasting experience (*FEXP*) are used to identify the 'worst' and 'best' quality analysts. In Column 1 of Panel A (unrestricted sample) and Panel B (matched sample), my test variable (*INDSP*) is significantly and negatively correlated with *DIFABSFE* ($\beta = -0.1440$, *p*-values < 0.001). In the regressions in which I use analysts' firm-specific experience to identify the 'worst' and 'best' quality analysts (Column 2), I also obtain negative and significant coefficients for my test variable (*INDSP*) (Panel A: $\beta = -0.0816$, $p < 0.01$; Panel B: $\beta = -0.1143$, $p < 0.01$). Taken together, these results support my prediction that industry specialist auditors reduce absolute forecast errors to a greater extent when the forecasts are made by less experienced analysts, which in turn reduces the difference in forecast errors between the 'worst' and 'best' quality analysts.

Table 8.6: Difference in Long-Horizon Forecast Errors between the 'Worst' and 'Best' Analysts against Audit Firm Industry Specialisation (Tests of H3b)

Panel A OLS Regressions on Unrestricted Samples

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	–0.1440***	–0.0816***	–0.0509	–0.0320	–0.1654***	–0.1621***
		(<0.001)	(0.009)	(0.101)	(0.117)	(<0.001)	(<0.001)
ABSFE_B	–	–8.4488***	–8.2844***	–10.4029***	–5.5898***	–10.2225***	–8.8123***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
DIFANQ	–	0.0055	0.0062	–0.0003		–0.0002	0.0037
		(0.309)	(0.562)	(0.628)		(0.972)	(0.656)
ANQ_B	+	–0.0058	–0.0070	0.0000		0.0025	–0.0025
		(0.294)	(0.503)	(0.992)		(0.587)	(0.723)
HORIZON_B	+	0.0008***	0.0012***	0.0007*	0.0006*	0.0008**	0.0009***
		(0.005)	(0.003)	(0.069)	(0.068)	(0.019)	(0.005)
DIFHORIZON	+	0.0038***	0.0041***	0.0042***	0.0031***	0.0040***	0.0042***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
DISP	+	3.2919	3.7148	12.1542***	6.4156***	7.3146	3.2008
		(0.242)	(0.182)	(<0.001)	(0.009)	(0.107)	(0.290)
SIZE	?	–0.0085	0.0030	0.0138	–0.0038	0.0022	0.0030
		(0.373)	(0.753)	(0.202)	(0.581)	(0.846)	(0.764)
NUMEST	–	0.0169	0.0181	–0.0020	0.0254*	–0.0454	0.0061
		(0.405)	(0.299)	(0.934)	(0.071)	(0.218)	(0.847)
ZSCORE	+	0.0096	0.0062	0.0233**	0.0208***	0.0291**	0.0024
		(0.354)	(0.518)	(0.034)	(0.007)	(0.023)	(0.827)
LOSS	+	0.2662***	0.2497***	0.2254***	0.1051*	0.3296***	0.2827***
		(<0.001)	(<0.001)	(<0.001)	(0.056)	(<0.001)	(<0.001)
ABSECHG	+	3.6513***	4.1667***	4.6941***	3.2101***	4.2578***	4.0463***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	–0.0839	–0.3203**	–0.2388	–0.1534	–0.2823*	–0.3388**
		(0.605)	(0.034)	(0.167)	(0.295)	(0.091)	(0.042)
EL	+	0.0184	0.0120	–0.0155	0.0038	0.0048	0.0096
		(0.109)	(0.282)	(0.204)	(0.634)	(0.695)	(0.385)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes	yes	yes
CONSTANT		0.0275	–0.2647	0.1955	–0.1620	–0.0625	–0.1582
		(0.930)	(0.191)	(0.429)	(0.173)	(0.749)	(0.498)
N		20,525	20,525	20,525	9,276	20,525	20,525
R^2		0.049	0.050	0.061	0.070	0.058	0.052

Robust p-values of the coefficients for *INDSP* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

Panel B OLS Regressions on Propensity Score Matched Samples

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.1362***	−0.1143***	−0.0652	−0.0542*	−0.1665***	−0.1866***
		(0.003)	(0.008)	(0.105)	(0.065)	(0.001)	(<0.001)
ABSFE_B	–	−7.5087***	−8.7732***	−7.9161***	−5.2555***	−8.7246***	−9.0078***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
DIFANQ	–	−0.0136	−0.0021	−0.0017		0.0141	0.0045
		(0.463)	(0.932)	(0.361)		(0.243)	(0.814)
ANQ_B	+	0.0181	0.0007	0.0000		−0.0117	−0.0035
		(0.333)	(0.976)	(0.999)		(0.277)	(0.831)
HORIZON_B	+	0.0003	0.0003	−0.0007	−0.0014	−0.0000	0.0001
		(0.559)	(0.613)	(0.168)	(0.169)	(0.988)	(0.905)
DIFHORIZON	+	0.0031***	0.0041***	0.0036***	0.0025**	0.0034***	0.0036***
		(<0.001)	(<0.001)	(<0.001)	(0.030)	(<0.001)	(<0.001)
DISP	+	12.2357*	−2.8012	21.0413**	1.6674	20.4877***	2.7630
		(0.062)	(0.700)	(0.014)	(0.831)	(0.008)	(0.697)
SIZE	?	0.0325	0.0362	0.0175	−0.0119	0.0330	0.0461*
		(0.191)	(0.137)	(0.534)	(0.459)	(0.278)	(0.077)
NUMEST	–	0.0529	0.0708	0.1391**	0.0731*	0.1230	0.0593
		(0.340)	(0.141)	(0.032)	(0.052)	(0.182)	(0.465)
ZSCORE	+	0.0104	−0.0018	0.0180	0.0343	0.0435	0.0072
		(0.708)	(0.944)	(0.566)	(0.106)	(0.168)	(0.797)
LOSS	+	0.3996***	0.4412***	0.3075**	0.1208	0.4232***	0.4587***
		(0.002)	(<0.001)	(0.037)	(0.445)	(0.004)	(<0.001)
ABSECHG	+	2.1453	4.8026***	2.8191*	4.4507***	3.1066*	3.6800**
		(0.128)	(0.001)	(0.060)	(0.002)	(0.067)	(0.015)
STDROE	+	0.2307	0.1582	−0.4747	−0.2150	−0.2520	0.2491
		(0.607)	(0.690)	(0.275)	(0.513)	(0.581)	(0.561)
EL	+	0.0055	0.0214	0.0092	0.0246	0.0278	0.0012
		(0.845)	(0.438)	(0.777)	(0.217)	(0.396)	(0.967)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes	yes	yes
CONSTANT		−0.2868	−0.3822	1.4156***	0.0436	−0.2391	−0.2299
		(0.340)	(0.231)	(<0.001)	(0.893)	(0.514)	(0.478)
N		3,312	3,312	3,312	1,462	3,312	3,312
R^2		0.060	0.072	0.064	0.074	0.076	0.069

Robust p-values of the coefficients for *INDSP* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

This table presents the results for the tests of H3b using OLS regressions on unrestricted samples (Panel A) and propensity score matched samples (Panel B). Column 1 of Panel A and B reports the results of the regressions using analysts' general forecasting experience (*GEXP*) to identify the 'worst' and 'best' quality analysts. Column 2 of Panel A and B reports the results of the regressions using analysts' firm-specific forecasting experience (*FEXP*) to identify the 'worst' and 'best' quality analysts. Column 3 of Panel A and B reports the results of the regressions using the size of analysts' employing brokerage firm (*BSIZE*) to identify the 'worst' and 'best' quality analysts. Column 4 of Panel A and B reports the results of the regressions using analysts' 'All-Star' status (*STAR*) to identify the 'worst' and 'best' quality analysts. Columns 5 and 6 of Panel A and B report the results of regressions using two composite score measures (*CSCORE3* and *CSCORE4*) to identify the 'worst' and 'best' quality analysts.

Variable Definitions: DIFABSFE is the absolute forecast error of the 'worst' quality analyst minus the absolute forecast error of the 'best' quality analyst where the 'worst' and 'best' quality analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4* where *GEXP* represents analysts' general experience, *FEXP* represents analysts' firm-specific experience, *BSIZE* represents the number of analysts employed by a brokerage firm, *STAR* equals 1 if an analyst was ranked by *II*'s All-America Research Team, 0 otherwise, *CSCORE3* is a composite measure of analyst quality, based on all four quality proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*), *CSCORE4* is the composite measure of analyst quality, based on three quality proxies representing the personal attributes of the analysts (*GEXP*, *FEXP* and *STAR*); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *ABSFE_B* is the absolute forecast error of the 'best' quality analyst, according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*; *ANQ_B* is the level of the quality proxy for the 'best' analyst, where the 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*; *DIFANQ* is the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4; HORIZON_B* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst, where 'best' analyst is determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*; *DIFHORIZON* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4; DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

Column 3 reports the results of the tests of difference in forecast accuracy according to analyst quality, where quality is proxied by the size of the broker employing the analysts (*BSIZE*). In both the unrestricted and matched samples, the coefficients for audit firm industry specialisation (*INDSP*) are negative but insignificant (Panel A: $p = 0.101$; Panel B: $p = 0.105$). This indicates that audit quality does not significantly reduce the difference in forecast accuracy between analysts employed by larger or smaller brokerage firms.

Column 4 reports the results for the tests of H3b using analysts' 'All-Star' status (*STAR*) to identify the 'worst' and 'best' quality analysts. The coefficient for my test variable (*INDSP*) is negative but insignificant in the unrestricted sample (Panel A).

However, it is significantly negative ($\beta = -0.0542$, $p = 0.065$) in the propensity score matched sample (Panel B). This supports my prediction that the hiring of a high-quality audit firm decreases the forecasting disadvantage of non-star analysts, consistent with the specialist auditors enhancing the quality of the published financial reports on which non-star analysts tend to more heavily rely.

In Columns 5 and 6, I report the results of the regressions of differences in absolute forecast errors across analysts of higher and lower quality, where quality is proxied by my two composite measures, *CSCORE3* and *CSCORE4*. Similar to the tests of H3a, *CSCORE3* is a function of all four analyst quality proxies (*GEXP*, *FEXP*, *BSIZE* and *STAR*), while *CSCORE4* only captures analysts' personal attributes (thus excluding *BSIZE*). The coefficients for audit firm industry specialisation (*INDSP*) are negative and highly significant (p-values <= 0.001) for both composite measures, regardless of whether the regressions are estimated on the unrestricted or propensity score matched samples. These results provide further support for my contention that audit quality improves the usefulness of published financial reports for predicting future performance and that such positive impact is greater for lower-quality analysts who rely more heavily on financial reports when generating their forecasts.

Since the models used to test H3b are exploratory, I discuss the statistical properties of the control variables employed in more detail than for earlier models. In addition to the firm-level control variables used in previous tests, I include five analyst-specific control variables for the tests of H3b, because the dependent variable in these tests is defined at the analyst level. The firm-level control variables are either insignificant or less significant in the tests of H3b than they are in the tests of the other hypotheses. While the control variables *ABSECHG* and *LOSS* have positive and significant coefficients consistently across different regressions, the other firm-year-level controls are not consistently significant across the various models presented in Panel A. Overall, the firm-level control variables have no consistent correlation with the difference in forecast accuracy between the 'worst' and 'best' analysts.

The coefficients for the analyst-related control variables contribute more greatly to model fit. For example, the coefficients for the level of the absolute forecast errors of the 'best' quality analyst (*ABSFE_B*), the forecast horizon of the 'best' analyst (*HORIZON_B*) and the difference in forecast horizon between the 'worst' and 'best' analysts (*DIFHORIZON*) are in the predicted directions and are significant in all unrestricted and matched samples. These results indicate that the timing of forecasts made by higher and lower-quality analysts and the pertaining underlying forecasting difficulty explain some of the difference in forecast accuracy across analysts. However, the variables capturing the extent to which the 'worst' and 'best' analysts differ in quality (*DIFANQ* and *ANQ_B*) are not correlated with the difference in their forecasting performance.

In summary, I find strong evidence to support H3b, which argues that audit firm industry specialisation reduces the difference in forecast accuracy of the 'worst' and 'best' analysts when these analysts are identified based on analysts' general experience, firm-specific experience or composite scores. I obtain similar evidence in a matched sample regression using the 'All-Star' analyst quality measure. Although H3b is not supported when brokerage size is exclusively used to proxy analyst quality, the partial impact of brokerage size on difference in forecast accuracy is implicitly captured through the composite score measure (*CSCORE3*), tests based on which support H3b. These results provide further support for my conjecture that the presence of an industry specialist auditor enhances the usefulness of published financial reports for predicting future performance, thereby improving the relative performance of lower analysts who I argue are relatively dependent on primary information sources such as financial reports when generating their forecasts. This evidence is again consistent with the existence of a causal relationship between auditor industry specialisation and the accuracy of analysts' prediction of earnings.

8.6 Sensitivity Tests for Hypotheses 3a and 3b

In this section, I conduct a range of sensitivity tests for both H3a and 3b, including variation in the lag-structure of controls (Section 8.6.1), and alternate specifications of the propensity score prediction regressions and the use of Heckman treatment effects regressions (Section 8.6.2). I also report robustness tests specific to each hypothesis in Section 8.6.3 (H3a) and Section 8.6.4 (H3b). I now describe these tests in turn.

8.6.1 Changes in Lag-Structure of Controls

The control variables (*ABSECHG*, *DISP* and *ZSCORE*) that I included in Models 5a, 5b and 6 may be affected by the realisation of the earnings being forecast, which is subject to the impact of auditor identity. Therefore, I re-estimated the models using the one-year lags of these controls to examine the robustness of the reported results. In these untabulated regressions, my main results are not substantively affected.

8.6.2 Alternate Specifications of the Two-stage Regressions

My main tests of H3a and H3b use a PSM approach to control for the endogenous selection of industry specialist auditors. In this section, I test the sensitivity of my PSM regression results to variations in the specification of the first-stage 'matching' equation, and to the use of Heckman treatment effects regressions. For brevity, these results are not tabulated but are discussed below.

I begin my sensitivity analyses by testing several alternate specifications of the first-stage matching regressions under the PSM approach. In the main tests, my first-stage regressions, which estimate the conditional probability of a client hiring an industry specialist, include only the second-stage controls as predictors, with one exception—that I do not include the level of the best analysts' absolute forecast errors (*ABSFE_B*) in the first-stage regressions for tests of H3b. I add the lag of the client's operating cycle (in tests of H3a and H3b), analyst quality main effects (in tests of H3a) and the level of the 'best' analysts' absolute forecast errors (in tests of H3b). I generate results of similar direction and significance to my main results, with the single exception being that the interaction between *INDSP* and *CSCORE2* in the tests of H3a is marginally insignificant ($p = 0.11$).

Further, to test the sensitivity of my main results to matching on post-treatment variables, I re-estimate Models 5a, 5b and 6 using first-stage regressions that include the lagged (rather than current) values of *ABSECHG*, *ZSCORE* and *DISP*, and predictors that are theoretically independent of the current-period auditor identity (e.g. replacing *ABSECHG* with the absolute change in clients' cash flow from operations, and replacing *STDROE* with the standard deviation of cash flow from operations). The overall model fit decreases for these untabulated regressions, but the coefficients for my test variables are in the predicted directions and of similar significance to the main results.

I next employ the Heckman treatment effects regressions to examine the robustness of the results generated by the PSM approach. I commence this analysis using the same instruments employed in the 2SLS regressions (*INDRELSIZE* and *CYCLE*). For the tests of H3a, I also interact these instruments (*INDRELSIZE* and *CYCLE*) with the analyst quality proxies and include these interaction terms as additional instrumental variables in the first-stage Heckman regressions. While I obtain coefficients similar to my main results, the Sargan tests suggest that my instruments are not exogenous (e.g. Sargan test: *CSCORE1* $p = 0.0127$; *CSCORE2* $p = 0.005$). I test alternate candidate instruments, such as clients' capital intensity, price-earnings ratio and stock issuance, but I cannot reject the possibility of the instruments' endogeneity in any of the tests. For the tests of H3b, the specification tests show that my instruments are both relevant and exogenous; thus, the resulting estimators are consistent. The coefficients for my test variable (*INDSP*) are significantly negative (*CSCORE3* $\beta = -0.0897$, $p = 0.017$; *CSCORE4* $\beta = -0.0542$, $p = 0.075$), consistent with my main findings that audit quality reduces the difference in forecast accuracy between the 'worst' and 'best' analysts.

In summary, my main results for the tests of H3a and H3b are robust to the various specifications of the matching regressions and are not tainted by post-treatment bias. The Heckman treatment-effect regressions confirm my findings for the tests of H3b, for which exogenous instruments can be identified and confirmed. For tests of H3a, the coefficients for the test variables are consistent with my main results, but these findings should be interpreted with caution because the specification tests fail to confirm the exogeneity of the instruments.

8.6.3 Other Sensitivity Tests for H3a

I conduct a range of additional tests of the robustness of my main results for tests of H3a, which I describe below. My main tests of H3a use the average quality of analysts following a client firm to proxy the overall quality of that cohort of analysts. However, the quality of the best analyst following a firm may be of greater empirical significance, because other analysts may use the expert's forecasts to inform their own (this phenomenon is known as 'herding'). Consequently, I re-estimate Models 5a and 5b using the maximum value of the analyst quality proxies within the cohort of analysts following a firm as my proxy for analyst quality. For brevity, in Table 8.7 I tabulate only the results for the regressions using the maximum value of the composite score measures. The coefficients for my test variables (interactions between *INDSP* and the relevant analyst quality proxies) remain positive and significant, providing further support for my prediction that the association between auditor industry specialisation and forecast accuracy decreases with analyst quality.

As my 'All-Star' proxy likely reflects the other analyst quality proxies, I exclude this proxy (*STAR*) in the calculation of the composite measures and re-estimate Models 5a and 5b using these composite measures. Once more, my untabulated test coefficients improve in their significance levels, or are substantively unaffected by this modelling choice.

Table 8.7: Tests of H3a using the Maximum Value of Analyst Quality Proxies

Dependent Variable: *ABSFE*					
Columns		(1)	(2)	(3)	(4)
		Unrestricted Samples		Matched Samples	
	Pred.	*CSCORE1M*	*CSCORE2M*	*CSCORE1M*	*CSCORE2M*
INDSP	–	–0.0037***	–0.0031**	–0.0092***	–0.0090***
		(0.003)	(0.011)	(0.003)	(0.003)
CSCOREM	–	–0.0002***	–0.0019***	–0.0003***	–0.0034***
		(0.003)	(0.004)	(<0.001)	(<0.001)
*INDSP*CSCOREM*	+	0.0000**	0.0005	0.0001***	0.0018***
		(0.041)	(0.112)	(0.006)	(0.005)
FFOLLOW	?	–0.0000	–0.0000	0.0000	–0.0000
		(0.514)	(0.466)	(1.000)	(0.989)
IFOLLOW	?	0.0004***	0.0004***	0.0003	0.0004
		(<0.001)	(<0.001)	(0.314)	(0.257)
DISP	+	0.3990***	0.3988***	0.2607	0.2609
		(<0.001)	(<0.001)	(0.231)	(0.230)
HORIZON	+	0.0205***	0.0207***	0.0235***	0.0236***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
SIZE	?	0.0005	0.0004	–0.0002	–0.0002
		(0.353)	(0.388)	(0.740)	(0.676)
NUMEST	–	–0.0005	–0.0005	0.0020	0.0019
		(0.623)	(0.607)	(0.188)	(0.231)
ZSCORE	+	0.0011*	0.0011*	0.0010	0.0010
		(0.055)	(0.057)	(0.222)	(0.237)
LOSS	+	0.0192***	0.0192***	0.0219***	0.0219***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.4951***	0.4952***	0.5755***	0.5753***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	–0.0224***	–0.0226***	–0.0256***	–0.0261***
		(<0.001)	(<0.001)	(0.003)	(0.003)
EL	+	–0.0013***	–0.0014***	–0.0013***	–0.0014***
		(0.001)	(0.001)	(0.010)	(0.006)
INDUSTRY		yes	yes	yes	yes
CONSTANT		–0.0908***	–0.0920***	–0.1156***	–0.1161***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
N		22,742	22,742	3,740	3,740
R^2		0.559	0.559	0.612	0.612

Robust p-values of the coefficients for *INDSP*, *CSCORE* and *INDSP*CSCORE* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1).

This table presents the results for the tests of H3a based on regressions using the maximum value of the four-quality composite measure (*CSCORE1M*) (Columns [1] and [3]), and the three-quality composite measure (*CSCORE2M*) of analyst quality on both the unrestricted and propensity score matched samples.

Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *CSCOREM* is a composite measure of analyst quality calculated based on the maximum value of the four singular analyst quality proxies (*CSCORE1M*), or the three quality proxies representing the personal attributes of the analysts (*CSCORE2M*); *INDSP*CSCORE* is the interaction between *INDSP* and *CSCORE1M* or *CSCORE2M*; *FFOLLOW* is the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *IFOLLOW* is the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

8.6.4 Other Sensitivity Tests for H3b

To further examine the robustness of my results for the tests of H3b, I conduct a range of tests. To eliminate the mechanical correlation between the forecasts made by the 'worst' and 'best' analysts and my control for observed dispersion (*DISP*), which has the potential to obscure the interpretation of coefficients, I re-compute *DISP* by (a) excluding the 'worst' and 'best' analysts' forecasts[121] and (b) replacing *DISP* with the forecast range, equal to the difference between the highest and lowest forecast earnings for a firm, excluding the forecasts of the 'worst' and 'best' analysts. For parsimony, I tabulate only the coefficients for my test variable (*INDSP*), in Table 8.8. In Panels A and B, I find that my main results are robust to variation in the measurement of *DISP* and the use of forecast range as a substitute for *DISP*.

My main results for the tests of H3b are based on models that include a control for the absolute forecast errors of the 'best' analyst. As this variable is likely to be structurally related to my treatment variable (*INDSP*), I re-estimate my regressions by (a) replacing this measure with the forecast error of the 'worst' analyst, (b) replacing this measure with the mean value of the forecast errors of the 'best' and 'worst' analysts and[122] (c) excluding this control altogether. I tabulate the results for the models using the forecast errors of the 'worst' analyst in Table 8.9. While some of these models have slightly higher R^2 statistics than those reported in Table 8.6, the coefficients for my test variables (*INDSP*) remain negative and significant at similar or stronger confidence levels.

---

[121] This computation of dispersion is not practical for the 'All-Star' model, as most of the firms included are followed by just one 'All-Star' analyst.

[122] In models in which I use the forecast accuracy of the 'worst' analyst or the average of the forecast accuracy of the 'best' and 'worst' analysts, I change the value of a quality proxy (*ANQ_B*) and the forecast horizon (*HORIZON_B*) accordingly.

Table 8.8: Tests of H3b using Alternative Controls for Forecast Dispersion

Panel A Re-computation of Forecast Dispersion (*DISP*)

Dependent Variable: *DIFABSFE*						
Columns		(1)	(2)	(3)	(4)	(5)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*CSCORE3*	*CSCORE4*
Unrestricted Sample						
INDSP	–	−0.1202***	−0.0659*	−0.0440	−0.1432***	−0.1562***
		(0.003)	(0.052)	(0.186)	(0.004)	(<0.001)
DISP_excl	+	3.7601	4.9536*	7.2149**	5.0986	3.4995
		(0.203)	(0.083)	(0.042)	(0.377)	(0.312)
Matched Sample						
INDSP	–	−0.1977***	−0.0816*	−0.0426	−0.1480**	−0.2227***
		(<0.001)	(0.06)	(0.250)	(0.015)	(<0.001)
DISP_excl	+	12.3264	3.4222	32.4227***	31.3349**	8.5522
		(0.147)	(0.679)	(<0.001)	(0.014)	(0.355)

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

Panel B Forecast Range as a Replacement for Forecast Dispersion (*DISP*)

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
Unrestricted Sample							
INDSP	–	−0.1437***	−0.0812***	−0.0499	−0.0328	−0.1648***	−0.1618***
		(<0.001)	(0.009)	(0.106)	(0.113)	(<0.001)	(<0.001)
RANGE	+	0.0038	−0.0015	0.0018	0.0038	0.0131	0.0079
		(0.579)	(0.728)	(0.750)	(0.606)	(0.171)	(0.406)
Matched Sample							
INDSP	–	−0.1963***	−0.1172***	−0.0536	−0.0609**	−0.1705***	−0.2119***
		(<0.001)	(0.005)	(0.155)	(0.047)	(<0.001)	(<0.001)
RANGE	+	0.0681	0.1283*	0.1953**	0.0240	0.3198**	0.0703
		(0.472)	(0.086)	(0.037)	(0.655)	(0.012)	(0.472)

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

This table presents the results for the tests of H3b using alternative controls for forecast dispersion. Panel A reports the results of the regressions using the re-computed forecast dispersion on both the unrestricted and propensity score matched samples. Panel B reports the results of the regressions using the forecast range as a replacement for dispersion on both the unrestricted and propensity score matched samples.

Variable Definitions: DIFABSFE is the difference in absolute forecast errors between the 'worst' and 'best' quality analysts, measured as per Equation (7); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *DISP_excl* is the forecast dispersion (*DISP*), excluding of the forecasts of the 'worst' and 'best' analysts; *RANGE* is the difference between the highest and lowest forecast earnings for a firm, excluding of the forecasts of the 'worst' and 'best' analysts.

Table 8.9: Tests of H3b using Alternate Controls for the Value of the 'Best' Quality Analyst

Panel A OLS Regressions on Unrestricted Samples

Dependent Variable: *DIFABSFE*

Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.1128***	−0.0506*	−0.0111	−0.0442*	−0.1231***	−0.1265***
		(<0.001)	(0.07)	(0.377)	(0.05)	(0.001)	(<0.001)
ABSFE_W	+	10.0033***	8.8276***	12.2283***	1.3733**	12.3598***	10.5114***
		(<0.001)	(<0.001)	(<0.001)	(0.016)	(<0.001)	(<0.001)
DIFANQ	–	−0.0039*	−0.0019	0.0000		0.0019	0.0007
		(0.081)	(0.370)	(0.959)		(0.148)	(0.745)
ANQ_W	–	−0.0061	−0.0066	−0.0001		−0.0035	−0.0067
		(0.274)	(0.531)	(0.920)		(0.455)	(0.338)
HORIZON_W	–	−0.0003	0.0000	−0.0007**	−0.0000	−0.0005	−0.0003
		(0.390)	(0.934)	(0.049)	(0.916)	(0.114)	(0.367)
DIFHORIZON	+	0.0035***	0.0035***	0.0040***	0.0028***	0.0039***	0.0040***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
N		20,525	20,525	20,525	9,276	20,525	20,525
R^2		0.065	0.056	0.080	0.028	0.081	0.070

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Coefficient estimates are multiplied by 100 in this table.

Panel B OLS Regressions on Propensity Score Matched Samples

Dependent Variable: *DIFABSFE*

Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.1025**	−0.0856**	−0.0227	−0.0652*	−0.1249**	−0.1552***
		(0.017)	(0.034)	(0.323)	(0.073)	(0.037)	(0.001)
ABSFE_W	+	9.9311***	7.1590***	13.2741***	3.0690**	12.7045***	9.2474***
		(<0.001)	(<0.001)	(<0.001)	(0.030)	(<0.001)	(<0.001)
DIFANQ	–	−0.0018	−0.0024	−0.0006		0.0024	0.0003
		(0.752)	(0.659)	(0.414)		(0.423)	(0.955)
ANQ_W	–	0.0151	0.0064	−0.0003		−0.0195*	−0.0072
		(0.417)	(0.795)	(0.870)		(0.064)	(0.652)
HORIZON_W	–	−0.0004	−0.0004	−0.0016*	−0.0020**	−0.0008	−0.0007
		(0.646)	(0.490)	(0.086)	(0.041)	(0.237)	(0.366)
DIFHORIZON	+	0.0031***	0.0041***	0.0042***	0.0042***	0.0040***	0.0039***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
N		3,312	3,312	3,312	1,462	3,312	3,312
R^2		0.085	0.056	0.118	0.053	0.116	0.071

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Coefficient estimates are multiplied by 100 in the table.

This table presents the results for the tests of H3b using the controls constructed based on the value of the 'worst' quality analyst (i.e. *ABSFE_W*, *HORIZON_W* and *ANQ_W*). Panel A (Panel B) reports the results of OLS regressions on unrestricted samples (propensity score matched samples).

Variable Definitions: *DIFABSFE* is the difference in absolute forecast errors between the 'worst' and 'best' quality analysts, measured as per Equation (7); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *ABSFE_W* is the absolute forecast error of the 'worst' quality analyst, according to various analyst quality proxies; *DIFANQ* is the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies; *ANQ_W* is the level of the quality proxy for the 'worst' analyst, where the 'worst' analyst is determined according to various analyst quality proxies; *HORIZON_W* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst, where 'worst' analyst is determined according to various analyst quality proxies; *DIFHORIZON* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies.

My main tests for H3b are based on the dichotomous measure of audit firm industry specialisation. I examine the impact of this modelling choice and re-estimate Model 6 using the continuous measure of industry specialisation. Table 8.10 tabulates the results for my test variable (*INDSP*) using both OLS (Panel A) and 2SLS regressions (Panel B). In regressions in which endogeneity exists (*BSIZE* and *STAR*), the 2SLS coefficient estimates show that auditor industry specialisation is significantly and negatively correlated with the difference in forecast accuracy, consistent with my main results. In regressions in which the DWH statistics show no evidence of endogeneity (*GEXP*, *FEXP*, *CSCORE3* and *CSCORE4*), the OLS estimates are efficient and consistent and the coefficients are significantly negative for models based on the *GEXP*, *CSCORE3* and *CSCORE4* analyst quality proxies. In summary, these results are generally consistent with my main results, which suggest that audit quality reduces the difference in forecast accuracy between the 'worst' and 'best' analysts. In fact, while I find no evidence that *INDSP* decreases *DIFABSFE* in models based on *BSIZE* in the main tests, my sensitivity analyses show that the continuous measure of auditor industry specialisation reduces the difference in forecast accuracy between analysts employed by smaller and larger brokerage firms.

Table 8.10: Tests of H3b using the Continuous Measure of Portfolio-Share Audit Firm Industry Specialisation

Panel A OLS Regressions

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.6375**	−0.3746	−0.2583	0.1142	−0.5866**	−0.7075**
		(0.036)	(0.139)	(0.230)	(0.303)	(0.049)	(0.024)
N		20,525	20,525	20,525	9,276	20,525	20,525
R^2		0.048	0.050	0.061	0.070	0.058	0.051

Robust *p*-values of the coefficients are one-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

Panel B 2SLS Regressions for Endogeneity Corrections

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.5137	−1.0988***	−1.3843***	−0.5875**	−1.1616**	−0.7022*
		(0.141)	(0.007)	(0.005)	(0.041)	(0.016)	(0.073)
N		19,402	19,402	19,402	8,628	19,402	19,402
R^2		0.042	0.045	0.053	0.051	0.051	0.044
Durbin-Wu-Hausman test (*p*-values)		0.9085	0.9085	0.2717	0.0865	0.0312	0.3697
Sargan test (*p*-values)		0.4703	0.4703	0.6217	0.4162	0.7396	0.2883
Partial-R^2 of instruments in first-stage		0.1971	0.1971	0.1966	0.1977	0.2061	0.1973
Wald F-statistics		271.204	271.594	271.023	224.652	272.692	269.345
LM statistics		909.034	919.446	906.987	472.995	914.585	911.558

Robust *p*-values of the coefficients are one-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

This table presents the results for the tests of H3b using the continuous measure of portfolio-share audit firm industry specialisation. Panel A reports the results of OLS regressions. Panel B reports the results of 2SLS regressions.

Variable Definitions: *DIFABSFE* is the difference in absolute forecast errors between the 'worst' and 'best' quality analysts, measured as per Equation (7); *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation, measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm.

To examine whether my results for the tests based on the composite measures are sensitive to the weight of the 'All-Star' rankings, I re-estimate Model 6 after re-computing the composite score measure (*CSCORE3*) by assigning 'All-Star' analysts a ranking value of 3, 5 or 10, while assigning non-star analysts a ranking value of 1, or excluding the 'All-Star' proxy in the calculation of the composite measures. This is because 'All-Star' analysts receive a much larger ranking value when a firm is followed by a relatively large number of predominantly non-star analysts, which results in a greater distance between the ranking of 'All-Star' and non-star analysts. For example, if a firm is followed by 21 analysts, the only 'All-Star' analyst in the firm would be ranked 21, while the remaining non-stars would be assigned the average value of 10.5. However, if a firm is followed by three analysts, the only 'All-Star' analyst has a ranking value of three, while the others have an average ranking value of 1.5. My tabulated main results remain qualitatively the same when I alter the weight of the rankings and exclude the 'All-Star' proxy in the calculation of the composite measures.

8.7 Additional Tests

I conduct two additional tests of Hypotheses 3a and 3b. These tests comprise the use of alternate industry specialisation measures and the examination of the impact of the Reg FD, which may shed light on the nature of analyst expertise and its association with audit quality.

8.7.1 Alternate Industry Specialisation Measures

My main findings for the tests of H3a and H3b suggest that auditor portfolio-share industry specialisation has a greater (lesser) impact on the forecast accuracy of lower (higher) quality analysts. Similar to the additional analyses applied to the tests of H1a, H1b and H2, I use the alternate industry specialisation measures described in Chapter 7 to examine whether my main results are affected. For brevity, these additional tests focus on the composite score measures of analyst quality (*CSCORE1* and *CSCORE2* for H3a, *CSCORE3* and *CSCORE4* for H3b).

Table 8.11 Panel A reports the propensity score matched sample results for the sensitivity tests of H3a. In regressions in which the national-level weighted market-share measure is used to estimate *INDSP*, the coefficients for the interactions (*INDSP*CSCORE1* $\beta = 0.0002$, $p = 0.032$; *INDSP*CSCORE2* $\beta = 0.0023$, $p = 0.07$) are significantly positive (as predicted), while the coefficients for the main effects are significant and of the predicted signs. However, I do not obtain significant coefficients for the interactions between the composite score measures and the other alternate measures of auditor industry specialisation. Panel B reports the matched sample results for the sensitivity tests of H3b using the alternate industry specialisation measures. Once more, my main results for the tests of H3b are reinforced when the national-level weighted market-share measure is used to estimate *INDSP*. I find some evidence in support of a negative association between the national- or city-level market-share-defined *INDSP* and the difference in forecast errors between the 'best' and 'worst' analysts, where analysts are identified based on the composite score measures; however, each of these results is specific to a particular composite score measure.

Table 8.11: Tests of H3a and H3b using Alternate Audit Firm Industry Specialisation Measures

Panel A Tests of H3a

Dependent Variable: *ABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
		National-level INDSP_market		*National-level INDSP_weighted_market*		*City-level INDSP_market*	
	Pred.	*CSCORE1*	*CSCORE2*	*CSCORE1*	*CSCORE2*	*CSCORE1*	*CSCORE2*
INDSP	–	0.0007	0.0002	−0.0097***	−0.0089**	−0.0008	0.0002
		(0.355)	(0.448)	(0.003)	(0.010)	(0.326)	(0.457)
CSCORE	–	−0.0021**	−0.0002*	−0.0003**	−0.0033**	−0.0001*	−0.0011
		(0.033)	(0.030)	(0.012)	(0.02)	(0.096)	(0.135)
*INDSP*CSCORE*	+	−0.0004	−0.0000	0.0002**	0.0023*	0.0000	−0.0002
		(0.247)	(0.334)	(0.032)	(0.07)	(0.427)	(0.373)
FFOLLOW	?	−0.0001	−0.0001	0.0000	0.0001	−0.0001	−0.0001
		(0.287)	(0.284)	(0.626)	(0.583)	(0.445)	(0.427)
IFOLLOW	?	0.0005***	0.0004***	0.0003	0.0004	0.0005***	0.0005***
		(0.001)	(0.003)	(0.406)	(0.360)	(0.008)	(0.006)
DISP	+	0.4374***	0.4379***	0.3514*	0.3507*	0.4257***	0.4258***
		(<0.001)	(<0.001)	(0.070)	(0.070)	(<0.001)	(<0.001)
HORIZON	+	0.0223***	0.0222***	0.0150***	0.0151**	0.0221***	0.0222***
		(<0.001)	(<0.001)	(0.009)	(0.010)	(0.004)	(0.004)
SIZE	?	0.0006	0.0007	0.0006	0.0005	0.0012	0.0012
		(0.190)	(0.153)	(0.544)	(0.594)	(0.217)	(0.228)
NUMEST	–	−0.0011*	−0.0011*	−0.0003	−0.0004	−0.0005	−0.0005
		(0.080)	(0.081)	(0.850)	(0.830)	(0.619)	(0.585)
ZSCORE	+	0.0012**	0.0012**	0.0007	0.0007	0.0007	0.0007
		(0.037)	(0.036)	(0.536)	(0.549)	(0.120)	(0.120)
LOSS	+	0.0195***	0.0194***	0.0220***	0.0221***	0.0149***	0.0150***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	+	0.4818***	0.4817***	0.5147***	0.5143***	0.4882***	0.4883***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	−0.0218***	−0.0218***	−0.0130	−0.0131	−0.0215***	−0.0216***
		(<0.001)	(<0.001)	(0.163)	(0.158)	(<0.001)	(<0.001)
EL	+	−0.0017***	−0.0017***	−0.0014	−0.0014	−0.0019***	−0.0019***
		(0.003)	(0.004)	(0.217)	(0.224)	(0.009)	(0.008)
INDUSTRY		yes	yes	yes	yes	yes	yes
CONSTANT		−0.1207***	−0.1201***	−0.0828**	−0.0831**	−0.1303***	−0.1312***
		(<0.001)	(<0.001)	(0.016)	(0.017)	(0.004)	(0.004)
N		14,098	14,098	3,694	3,694	5,020	5,020
R^2		0.552	0.551	0.554	0.554	0.552	0.552

Robust *p*-values of the coefficients for *INDSP*, *CSCORE* and *INDSP*CSCORE* are one-tailed reported, and for other variables are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the propensity score matched sample regression results for the tests of H3a using alternate audit firm industry specialisation measures across models based on the composite measure of the analyst quality proxies (Model 5b). Columns (1) and (2) report the results of the regressions using the national-level market-share measure of industry specialisation. Columns (3) and (4) report the results of the regressions using the national-level weighted market-share measure of industry specialisation. Columns (5) and (6) report the results of the regressions using the city-level market-share measure of industry specialisation.

Panel B Tests of H3b

Dependent Variable: *DIFABSFE* Columns		(1)	(2)	(3)	(4)	(5)	(6)
		National-Level INDSP_market		*National-Level INDSP_weighted_market*		*City-Level INDSP_market*	
	Pred.	*CSCORE3*	*CSCORE4*	*CSCORE3*	*CSCORE4*	*CSCORE3*	*CSCORE4*
INDSP	–	−0.0561**	−0.0247	−0.1126*	−0.1310**	0.0276	−0.0586*
		(0.016)	(0.158)	(0.052)	(0.023)	(0.272)	(0.072)
ABSFE_B	–	−10.4061***	−9.0877***	−8.7006***	−8.9900***	−12.5878***	−9.5852***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
DIFANQ	–	0.0078	0.0071	0.0142	0.0051	0.0033	−0.0108
		(0.202)	(0.503)	(0.238)	(0.789)	(0.754)	(0.462)
ANQ_B	+	−0.0050	−0.0058	−0.0118	−0.0040	−0.0016	0.0065
		(0.363)	(0.522)	(0.270)	(0.803)	(0.862)	(0.603)
HORIZON_B	+	0.0007	0.0012**	−0.0001	0.0000	−0.0000	0.0008
		(0.140)	(0.017)	(0.923)	(0.964)	(0.981)	(0.397)
DIFHORIZON	+	0.0038***	0.0040***	0.0033***	0.0036***	0.0039***	0.0047***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)
DISP	+	7.7487	1.2093	20.5156***	2.7996	−5.6715	−7.7168
		(0.161)	(0.778)	(0.008)	(0.696)	(0.602)	(0.115)
SIZE	?	−0.0030	−0.0023	0.0348	0.0481*	0.0160	0.0173
		(0.840)	(0.859)	(0.257)	(0.068)	(0.486)	(0.405)
NUMEST	–	0.0009	0.0285	0.1239	0.0613	−0.0405	0.0605
		(0.984)	(0.477)	(0.179)	(0.451)	(0.593)	(0.394)
ZSCORE	+	0.0202	0.0016	0.0429	0.0064	0.0154	−0.0121
		(0.217)	(0.911)	(0.174)	(0.819)	(0.565)	(0.571)
LOSS	+	0.2492***	0.2238***	0.4355***	0.4730***	0.4581***	0.1715
		(0.003)	(0.003)	(0.003)	(<0.001)	(0.001)	(0.123)
ABSECHG	+	3.5125***	3.5786***	3.0719*	3.6451**	4.5357**	4.7552***
		(<0.001)	(<0.001)	(0.071)	(0.016)	(0.021)	(<0.001)
STDROE	+	−0.2863	−0.3659*	−0.2405	0.2619	−0.1573	−0.3314
		(0.189)	(0.094)	(0.599)	(0.542)	(0.674)	(0.336)
EL	+	−0.0028	−0.0038	0.0298	0.0034	−0.0075	−0.0347
		(0.868)	(0.797)	(0.364)	(0.904)	(0.769)	(0.121)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes	yes	yes

CONSTANT	0.4012	0.3035	−0.4224	−0.4398	0.1052	−0.7409**
	(0.317)	(0.525)	(0.255)	(0.177)	(0.774)	(0.039)
N	12,808	12,808	3,312	3,312	4,576	4,576
R^2	0.060	0.053	0.074	0.066	0.094	0.085

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$). Coefficient estimates are multiplied by 100 in this table.

This table presents the propensity score matched sample regression results for the tests of H3b using alternate audit firm industry specialisation measures across models based on composite measure of the analyst quality proxies (Model 6). Columns (1) and (2) report the results of regressions using the national-level market-share measure of industry specialisation. Columns (3) and (4) report the results of regressions using the national-level weighted market-share measure of industry specialisation. Columns (5) and (6) report the results of regressions using the city-level market-share measure of industry specialisation.

Variable Definitions: *DIFABSFE* is the difference in absolute forecast errors between the 'worst' and 'best' quality analysts, measured as per Equation (7); *INDSP* equals 1 if the client is audited by a national industry specialist who has the largest market share in a two-digit SIC industry (National-level *INDSP*_market), where the market share is calculated as per Equation (11), 0 otherwise; or, if the client is audited by a national industry specialist who has the largest weighted market share in a two-digit SIC industry (National-level *INDSP*_weighted_market), where the weighted market share is calculated as per Equation (12), 0 otherwise; or, if the client is audited by a city-industry specialist who has the largest market share in a two-digit SIC industry in a city defined as the U.S. Census Bureau definition of Metropolitan Statistical Area (City-level *INDSP*_market), where the city market share is calculated as per Equation (13), 0 otherwise; *CSCORE* is a composite measure of analyst quality, based on either all four analyst quality proxies (*CSCORE1*) or the three quality proxies representing the personal attributes of the analysts (*CSCORE2*); *INDSP**CSCORE* is the interaction between *INDSP* and *CSCORE1* or *CSCORE2*; *FFOLLOW* is the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; and; *IFOLLOW* is the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *ABSFE_B* is the absolute forecast error of the 'best' quality analyst, according to various analyst quality proxies: *GEXP*, *FEXP*, *BSIZE*, *STAR*, *CSCORE3* and *CSCORE4*; *ANQ_B* is the level of the quality proxy for the 'best' analyst, where the 'best' analyst is determined according to various analyst quality proxies; *DIFANQ* is the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies; *HORIZON_B* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst, where 'best' analyst is determined according to various analyst quality proxies; *DIFHORIZON* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

In summary, the weighted market-share measure of industry specialisation is associated with improved forecast accuracy, with this association being increased when the quality of analysts is lower. While I obtain mixed evidence concerning the relationship between market-share-defined industry specialisation measures, analyst quality and forecast accuracy, I am not surprised that the results for measures based on market share are weaker than are those for the portfolio-share metrics. This is because the market-share measure is strongly correlated with market dominance, a condition that economics has long recognised as having potential dysfunctional consequences. Moreover, Minutti-Meza (2013) has recently demonstrated the weakness of this proxy as a measure of audit quality.

8.7.2 Regulation Fair Disclosure

My primary results for the tests of H3a and H3b suggest that analysts with greater general or firm-specific experience, employed by a larger broker, designated as 'All-Star' or with higher composite ranking scores have a better understanding of the implications of earnings and rely less heavily on the quality of published financial reports (and thus audit quality) to predict earnings accurately. I argue earlier that this difference in forecasting performance reflects these analysts' greater access to information beyond that contained in the financial statements and their superior ability to process complex information.

However, analysts' access to private information is likely to have been reduced following the passage of Reg FD by the U.S. Securities and Exchange Commission (SEC), which became effective on 23 October 2000. Reg FD prohibits management's selective disclosure of material information to analysts and other investment professionals without disclosing the same information to the public. Although analysts' private information is not limited to that obtained from management,[123] the reduction in analysts' access to managerial information because of the regulation may result in a deterioration in the forecasting performance in the post-Reg FD period, with those analysts with superior pre-Reg FD performance likely to have been more greatly affected. Findlay and Mathew (2006) show that average analyst forecast accuracy decreases significantly after the passage of Reg FD. Similarly, others studies find that the positive association between analysts' general and firm-specific experience, brokerage size, 'All-Star' status and forecasting performance (i.e. forecast accuracy and the reduced accrual-related over-optimism) observed in the pre-Reg FD period is weaker or non-existent in the post-Reg FD period (Drake and Myers 2010; Keskek et al. 2013). Thus, to the extent that my singular and composite analyst quality proxies reflect the effects of access to managerial information, a weaker relation between analyst quality and forecast accuracy should be observed after the implementation of Reg FD. High-quality analysts should also be more dependent on the quality of financial reports for forecasting earnings post-Reg FD. As such, audit firm industry specialisation should have an increased impact on the forecast accuracy of these high-quality analysts in the post-Reg FD period. Of my proxies, the experience proxies are the most likely to be closely related to access to managerial information. This is because the longer an analyst follows a firm, the greater the likelihood that this analyst will develop a good relationship with management and gain access to private information. Therefore, I conjecture that the impact of Reg FD has a greater impact on my main results for tests employing the experience proxies.

---

[123] As noted earlier, analysts may have various sources of private information, such as from industry advocacy bodies, economic consultancies and firm management.

To examine the interaction associations among analyst quality, audit firm industry specialisation and forecast accuracy and Reg FD, I follow Drake and Myers (2010) to include an indicator variable (*REGFD*) that equals 1 in the post-Reg FD period (2002–2010) and 0 otherwise (1989–1999), [124] and interact this indicator variable with each of the test variables included in Models 5a, 5b and 6.

For parsimony, I tabulate only the key coefficients for my pre-post-Reg FD tests. Table 8.12 presents these coefficients for the tests of H3a (Panel A) and H3b (Panel B). For the tests of H3a (in Panel A), the coefficients for *INDSP* are significantly negative, while the untabulated sums of coefficients for *INDSP* and *INDSP*REGFD* are also negative and significant (Model5a: β = –0.0111, two-tailed $p < 0.01$; Model5b: *CSCORE1* β = –0.0086, two-tailed $p = 0.023$; *CSCORE2* β = –0.0063, two-tailed $p = 0.067$). These results suggest that audit firm industry specialisation has a negative effect on analysts' absolute forecast errors in both the pre- and post-Reg FD periods, consistent with my main results for the entire sample period. The coefficients for the interactions between *INDSP* and *GEXP, STAR, CSCORE1* and *CSCORE2* are positive and significant as expected; however, the coefficient for the interaction term *INDSP*GEXP*REGFD* is negative and significant. These results show that the moderating impact of analysts' general experience on the association between audit firm industry specialisation and forecast accuracy existing prior to Reg FD is less significant in the post-Reg FD period. In other words, analysts with a greater general experience may increase their dependence on the quality of published financial reports (and thus audit quality) in predicting client firms' future

---

[124] I exclude the years 2000–2001 to eliminate the possibility that analysts forecasting firms' 2000 and 2001 earnings benefitted from any managerial private information obtained prior to the enactment of Reg FD. Untabulated regressions based on a sample of 2001–2010 show similar results to those reported.

performance after the implementation of Reg FD. These results indicate that Reg FD is an exogenous shock from which I can test predictable changes in the test coefficients and establish evidence consistent with the importance of audit quality on analysts' forecasting performance.

Table 8.13 reports selected coefficients for the Reg FD tests of H3b. Consistent with my main findings for the entire sample period, the coefficients for *INDSP* and the sum of coefficients (untabulated) for *INDSP* and *INDSP*REGFD* are negative and significant in models where *GEXP*, *FEXP*, *CSCORE3* and *CSCORE4* are used to identify the 'worst' and 'best' analysts. This indicates that audit quality reduces the difference in forecast accuracy in both the pre- and post-Reg FD periods. While the coefficient for the interaction term (*INDSP*REGFD*) is insignificant in other models, it is positive and significant (β = 0.2519, two-tailed p = 0.041) in the *FEXP* model. These results suggest that Reg FD reduces the negative association between audit quality and the difference in forecast accuracy between analysts with the least and most firm-specific experience. These results are consistent with my conjecture that the improvements in the quality of financial reports driven by industry specialist auditors affect the forecasting performance of analyst with least firm-specific experience and analyst with most experience to the same extent in the post-Reg FD period because the latter one may benefit less from managerial information and rely more heavily on the quality of financial reports to predict earnings.

Table 8.12: Reg FD Tests of H3a

Dependent Variable: *ABSFE*				
Columns		(1)	(2)	(3)
	Pred.	Singular Proxy	*CSCORE1*	*CSCORE2*
REGFD	?	−0.0072***	−0.0067***	−0.0069***
		(<0.001)	(<0.001)	(<0.001)
INDSP	–	−0.0061*	−0.0074	−0.0082*
		(0.070)	(0.131)	(0.088)
*INDSP*REGFD*	–	−0.0050	−0.0012	0.0019
		(0.178)	(0.824)	(0.717)
GEXP	–	−0.0000		
		(0.813)		
*INDSP*GEXP*	+	0.0010**		
		(0.016)		
*INDSP*GEXP*REGFD*	?	−0.0007*		
		(0.073)		
FEXP	–	−0.0002		
		(0.136)		
*INDSP*FEXP*	+	0.0003		
		(0.592)		
*INDSP*FEXP*REGFD*	?	0.0004		
		(0.519)		
BSIZE	–	0.0002		
		(0.263)		
*INDSP*BSIZE*	+	−0.0005**		
		(0.027)		
*INDSP*BSIZE*REGFD*	?	0.0011***		
		(<0.001)		
STAR	–	−0.0037		
		(0.144)		
*INDSP*STAR*	+	0.0094**		
		(0.035)		
*INDSP*STAR*REGFD*	?	−0.0017		
		(0.642)		
CSCORE	–		−0.0000	0.0000
			(0.996)	(0.963)
*INDSP*CSCORE*	+		0.0039	0.0057*
			(0.117)	(0.075)
*INDSP*CSCORE*REGFD*	?		−0.0023	−0.0044
			(0.336)	(0.157)
N		20,735	20,735	20,735
R^2		0.553	0.552	0.552

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the Reg FD tests of H3a. Column 1 reports (Columns 2 and 3 report) the results of the regressions using the singular (two composite measures of) analyst quality proxies.

Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *REGFD* is an indicator variable which equals 1 in the post-Reg FD period (2002–2010) and 0 otherwise (1989–1999); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *GEXP* is the average general experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where general experience is measured as the number of years through year *t* for which an analyst *i* supplied at least one forecast for any firm; *FEXP* is the average firm experience of all analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where firm experience is measured as number of years through year *t* for which an analyst *i* supplied at least one forecast for firm *j*; *BSIZE* is the average brokerage size that employs analysts following a firm during the window over which long-horizon forecast accuracy is calculated, where brokerage size is measured as number of analysts employed by a broker employing analyst *i* who follows firm *j* in year *t*; *STAR* is the proportion of the analysts following firm *j*, during the long-horizon forecast window, who are ranked as an 'All-Star' by *II's* All-America Research Team in year *t*; *CSCORE* is a composite measure of analyst quality, based on either all four analyst quality proxies (*CSCORE1*) or the three quality proxies (*CSCORE2*); *INDSP*REGFD* is the interaction between *INDSP* and *REGFD*; *INDSP*GEXP* is the interaction between *INDSP* and *GEXP*; *INDSP*FEXP* is the interaction between *INDSP* and *FEXP*; *INDSP*BSIZE* is the interaction between *INDSP* and *BSIZE*; *INDSP*STAR* is the interaction between *INDSP* and *STAR*; *INDSP*CSCORE* is the interaction between *INDSP* and *CSCORE1* or *CSCORE2*; *INDSP*GEXP*REGFD* is the interaction between *INDSP*, *GEXP* and *REGFD*; *INDSP*FEXP* is the interaction between *INDSP*, *FEXP* and *REGFD*; *INDSP*BSIZE* is the interaction between *INDSP*, *BSIZE* and *REGFD*; *INDSP*STAR* is the interaction between *INDSP*, *STAR* and *REGFD*; *INDSP*CSCORE*REGFD* is the interaction between *INDSP*, *CSCORE1* (*CSCORE2*)and *REGFD*.

Table 8.13: Reg FD Tests of H3b

Dependent Variable: *DIFABSFE*							
Columns		(1)	(2)	(3)	(4)	(5)	(6)
	Pred.	*GEXP*	*FEXP*	*BSIZE*	*STAR*	*CSCORE3*	*CSCORE4*
REGFD	?	−0.0395	−0.0131	−0.2015***	−0.0332	−0.0312	0.0060
		(0.472)	(0.806)	(0.001)	(0.418)	(0.602)	(0.917)
INDSP	−	−0.1532***	−0.0914**	−0.0485	−0.0225	−0.1321***	−0.1586***
		(0.001)	(0.012)	(0.299)	(0.509)	(0.006)	(0.001)
*INDSP*REGFD*	?	0.0162	0.2519**	0.0285	−0.0305	−0.0540	−0.0213
		(0.714)	(0.041)	(0.549)	(0.374)	(0.260)	(0.643)
N		18,752	18,752	18,752	8,410	18,752	18,752
R^2		0.051	0.053	0.066	0.072	0.060	0.053

Robust p-values of the coefficients are two-tailed reported in parentheses (*** p<0.01, ** p<0.05, * p<0.1). Coefficient estimates are multiplied by 100 in this table.

This table presents the results for the Reg FD tests of H3b across all models.
Variable Definitions: *DIFABSFE* is the difference in absolute forecast errors between the 'worst' and 'best' quality analysts, measured as per Equation (7); *REGFD* is an indicator variable which equals 1 in the post-Reg FD period (2002–2010) and 0 otherwise (1989–1999); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *INDSP*REGFD* is the interaction between *INDSP* and *REGFD*.

In summary, I find limited evidence that Reg FD eliminates the moderating impact of average analyst quality on the association between audit firm industry specialisation and forecast accuracy. While there is evidence suggesting that a cohort of analysts with relatively greater general experience performs better and is less reliant on the quality of financial reports in predicting earnings prior to Reg FD, this superior performance is perhaps attributable to experienced analysts' greater access to managerial private information. When examining how audit firm industry specialisation affects the difference in forecast accuracy of the *individual* analysts following a client firm in the pre- and post-Reg FD periods, I find only an increase in the forecast accuracy of more (firm-specific) experienced analysts post-Reg FD. This is consistent with these experienced analysts' increased reliance on the quality of financials reports in forecasting earnings post-Reg FD. Overall, these results provide evidence consistent with the interpretation that high-quality analysts are less dependent on the quality of public information in generating forecasts, and that their superior performance is related to both their greater access to private information and their abilities in processing information. In the post-Reg period, this ability effect is likely to be more dominant.

8.8 Chapter Summary

In this chapter, I reported and discussed the results for the tests of H2, H3a and H3b. I presented evidence to support H2, which hypothesises that the negative impact of audit firm industry specialisation on the absolute forecast errors in analysts' prediction of earnings is more pronounced when client firm's operating risk is higher, consistent with a greater return to audit quality where the auditor's task complexity is higher. My results for the tests of H3a and H3b are generally consistent with my expectations and suggest that audit firm industry specialisation has a greater impact in improving analyst forecast accuracy when the quality of analysts covering a firm is lower. Overall, these results bring further convincing evidence to indicate that auditor industry specialisation causally affects analyst forecast accuracy.

CHAPTER 9: CONCLUSION

9.1 Introduction

This chapter concludes my thesis by summarising the findings (Section 9.2), discussing the implications of my thesis (Section 9.3), identifying and discussing its limitations (Section 9.4) and finally identifying potential opportunities for future research (Section 9.5).

9.2 Summary of Research Design and Findings

My thesis proposes and examines two research questions:

RQ 1: Is audit firm industry specialisation associated with analyst forecast accuracy?

RQ2: Does the strength of the association between audit firm specialisation and analyst forecast accuracy vary with factors affecting the relative importance of audit quality to the predictability of earnings?

Below, I describe how these research questions are addressed in my thesis and summarise the main findings.

9.2.1 Research Question One

My first research question examines whether audit firm industry specialisation is associated with analyst forecast accuracy. I conceive audit firm industry specialisation as the extent to which audit firms concentrate their productive activities in particular industries. I reviewed the relevant audit quality literature and developed a framework built thereon, to articulate the means by which the supply of high-quality audit services may improve client financial reporting quality, a measure of which is the predictability of earnings by sophisticated users (analysts). I argue that industry specialist auditors develop industry-specific expertise arising from these firms' strategic concentration of their services within specific industries. I further argue that this improves the auditors' judgment and decision making applied in client negotiations and other stages of the audit processes, leading to superior audit outcomes (such as improved earnings predictability). Based on this logic, two hypotheses are developed concerning the general association between audit firm industry specialisation and analyst forecast accuracy. Hypothesis 1a is a maintained hypothesis, the tests of which largely seek to reconcile prior findings. It predicts a non-directional association between analysts' short-horizon absolute forecast errors and audit firm industry specialisation. Hypothesis 1b predicts that analysts' long-horizon absolute forecast errors decrease with audit firm industry specialisation.

To test these hypotheses, I regress the absolute value of short- or long-horizon forecast errors (inverse measures of forecast accuracy) on audit firm portfolio-share industry specialisation (as a proxy for the supply of high-quality audit services) and a vector of control variables. My results for tests of H1a show that auditor industry specialisation does not consistently improve or impair analyst short-horizon forecast accuracy, and that the direction and significance of the measured relationship between these variables is highly susceptible to subtle changes in model specification. These results reflect the conflicting effects of audit quality on short-horizon forecast accuracy. On the one hand, industry specialist auditors improve the usefulness of financial reports for predicting future earnings; but at the same time, these specialist auditors are more likely to constrain clients' attempts to manipulate earnings to bias earnings towards market expectations. Conversely, my results for tests of H1b demonstrate that analyst long-horizon forecast accuracy increases with audit firm industry specialisation, and are robust to numerous model specifications and modelling choices. These results strongly support my predictions that earnings reports audited by an industry specialist auditor are more useful for predicting future earnings, thus increasing analyst forecast accuracy.

9.2.2 Research Question Two

My second research question concerns cross-sectional variation in the relationship between audit firm industry specialisation and forecast accuracy. The assumed sources of this variation are the underlying riskiness of client operations and the quality of the analysts covering a client firm. I argue that the impact of audit quality on the importance of published financial reports for predicting future earnings should be greater where the complexity of audit and forecasting tasks are greater and where these reports represent a relatively large proportion of the information used in forecasting. Consequently, I predict that audit firm industry specialisation has a greater impact in improving forecast accuracy when client firm's operating risk is higher (Hypothesis 2) and when the quality of analysts covering the client firm is lower (Hypotheses 3a and 3b).

To test Hypothesis 2, I proxy client firm's operating risk using cash flow volatility and include this measure and its interaction with auditor industry specialisation in a regression of analysts' absolute forecast errors. My results for tests of H2 indicate that the relationship between audit firm industry specialisation and forecast accuracy is more pronounced when client's operating risk is higher. This is consistent with the argument that the impact of audit quality increases with the difficulty of audit and forecasting tasks, and thereby suggesting greater confidence that the results of the tests of H1b are not purely driven by omitted variables.

To address the extent to which industry specialist auditors affect the relative forecast accuracy of high-quality and low-quality analysts (H3a and H3b), I draw on prior literature and measure analyst quality using analysts' general experience, firm-specific experience, brokerage size and 'All-Star' status. I also develop two composite measures to capture all aspects of analyst quality or only the analysts' personal attributes. To test H3a, which predicts that auditor industry specialisation is more strongly associated with forecast accuracy when the average quality of analysts is lower, I modify the base model to include proxies for the average quality of analysts following a client and their interactions with audit firm industry specialisation. To test H3b, which concerns the differential effect of audit quality on forecast accuracy across individual analysts, I regress the difference in forecast errors between the 'worst' and 'best' quality analysts against audit firm industry specialisation and controls. My results broadly support my predictions that the relationship between auditor industry specialisation and forecast accuracy is stronger when analyst quality is lower. This is consistent with audit quality having a greater impact on the usefulness of financial reports for predicting future performance for lower-quality analysts, who are expected to be more reliant on the quality of audited financial reports when issuing forecasts. Overall, these findings provide further evidence consistent with the existence of a causal, rather than spurious, relationship between audit quality and the usefulness of published earnings for predicting future performance.

9.3 Implications

The results reported and discussed in my thesis have several implications, including for regulators, scholars and investors. These implications are discussed in turn in the following sections.

9.3.1 Implications for Regulators

The FASB Conceptual Framework states that the objective of general-purpose financial reporting is to provide financial information relevant to decision making that will meet the needs of the maximum number of primary users, including existing and potential investors, lenders and other creditors (SFAC No. 8, FASB 2010 OB1; OB2; OB8 and BC1.9). Information relevant to the prediction of future earnings (OB.17) and cash flows (OB.18) is central to this objective. My thesis presents evidence that the forward-looking and 'user coverage' objectives of financial reporting are more likely to be fulfilled when audit quality is relatively high; a finding that should be of clear interest to regulators of financial reporting. In particular, my results suggest that the accuracy of analysts' prediction of clients' earnings is greater when the clients are audited by industry specialist auditors, which is consistent with superior quality audit services improving the quality of published general-purpose financial reports for predicting future performance (as per SFAC No. 8, FASB 2010 OB2, OB3 and BC1.31). I further demonstrate that the positive impact of audit quality on the usefulness of financial reports for predicting future performance is greater for those financial statement users who rely relatively heavily on the published financial reports in making decisions (i.e. analysts of lower quality). This indicates that higher-quality financial reports are of superior usefulness to a potentially greater number of financially competent users.

9.3.2 Implications for Scholars

My research also has implications for scholars. First, I reconcile and explain the inconsistent findings in the prior literature regarding the impact of audit firm industry specialisation on analyst forecast accuracy. In particular, I show that short-horizon regressions are highly sensitive to model specification, which is attributable to analyst short-horizon accuracy being confounded by the competing impacts of audit quality on the usefulness of historic reports and the likelihood of constraining benchmark-beating earnings management. I further demonstrate that long-horizon regressions consistently produce evidence that forecast accuracy increases with audit quality, supporting my argument that long-horizon forecasts are directly related to the quality of the audited reports and are less likely to be used as a benchmark towards which management manipulates future earnings to 'meet or just beat' analysts' forecasts. My thesis concludes that long-horizon forecast accuracy is a more direct and less noisy measure than short-horizon forecast accuracy, of the extent to which the objectives of financial reporting are satisfied.

Second, I present evidence that audit firm industry specialisation is associated with the accuracy of analysts' forecasts, and that this relationship is stronger in cases where theory predicts that audit quality should be of greater importance in explaining forecast accuracy. To this end, I show that auditor industry specialisation is more strongly associated with forecast accuracy when client firm's operating risk is higher and when the quality of the analysts following a firm is lower. These findings are of clear relevance to the auditing literature, which seeks to understand the economic role of audit quality and the factors that drive the provision of high-quality audit services.

Finally, my thesis addresses the current concerns regarding the endogenous selection of auditors in the auditor industry specialisation literature by identifying well-specified endogeneity models. In particular, prior studies fail either to account explicitly for the endogenous determination of auditor industry specialisation (BCK 2008; Payne 2008) or to use propensity score matched sample regressions to correct for the endogenous determination of auditor market-share industry specialisation (Minutti-Meza 2013). My study is the first to employ and validate a range of endogeneity-corrected regressions, including 2SLS, PSM and Heckman treatment-effect regressions, to control for both the dichotomous and continuous endogenous regressors (auditor portfolio-share industry specialisation). The methodologies applied here should be useful to scholars wishing to disentangle client characteristics from audit quality effects in their future research.

9.3.3 Implications for Investors

My thesis has implications for investors. My results show that analysts' predictions of earnings is more accurate for client firms audited by industry specialist auditors. These findings should be useful to investors for at least two reasons. First, analysts' earnings forecasts have been demonstrated as proxies for market expectations of future earnings (Kothari 2001; Affleck-Graves et al. 2002; Crabtree and Maher 2005; Dichev and Tang 2009). My study identifies an important cross-sectional determinant of analyst forecast accuracy, which indicates that the reliability of this proxy for market earnings expectations should be greater when audit quality is higher. Second, identifying conditions under which published earnings are more predictable are of clear interest to investors relying on analysts' forecasts as a basis for their resource allocation decisions.

9.4 Limitations

My thesis is subject to a number of limitations. While I do not consider that these limitations significantly bias the reported results, they should be taken into consideration when interpreting my findings.

First, empirical measures of audit firm industry specialisation are a noisy signal of the economic construct that they purport to measure, and thresholds used in the dichotomous quality measures are arbitrary. While I tested and demonstrated the robustness of my results using alternative thresholds to determine a dichotomous portfolio industry specialist, it remains possible that the results would be weaker if other arbitrary thresholds were employed. In addition, my sensitivity analyses show a less consistent and weaker association between auditor industry specialisation and forecast accuracy when alternate measures of specialisation are employed, each of which relies on 'market-share' metrics. However, these results are not surprising because the market-share measure is strongly correlated with market dominance, which may reduce competition and lead to negative audit outcomes. A recent study has demonstrated the weakness of this proxy as a measure of auditor industry specialisation (Minutti-Meza 2013). More importantly, as my thesis focuses on whether the extent to which audit firms concentrate their business in particular industries affects the quality of clients' published financial reports for analysts' predictions of future earnings, the portfolio-share measure of industry specialisation is of more direct relevance to the examination of this question.

Second, because the selection of an industry specialist auditor is not random, I control for endogeneity using various two-stage regressions, including 2SLS, PSM and Heckman regressions. Each approach has limitations, and while I have attempted to control for these limitations, such as through tests using alternate instruments for the 2SLS and Heckman regressions and using alternate first-stage matching variables for the PSM regressions, it remains possible that audit firm industry specialisation proxies some unobserved risk factors naturally correlated with the difficulty in forecasting earnings.

A further limitation is that the quality of analysts following a given firm may not be random, and this may affect my results, particularly those regarding the tests of Hypothesis 3a, which examines the average quality of analysts following a client firm.

Finally, my findings pertain directly to industry specialisation within U.S. Big N firms and may not be applicable outside this sample. My results may have limited implications in an international context in which the institutional settings and analyst behaviour are different to the U.S. context. Further, because I constrain my sample to clients of Big N firms, whether industry specialisation within non-Big N firms improves the usefulness of financial reports for predicting future performance is unknown. However, BCK (2008) and my untabulated statistics show that approximately 90 per cent of the U.S. public firms followed by at least three analysts are audited by a Big N audit firm, suggesting that my results are at the very least generalisable to a majority of U.S. public firms covered by financial analysts.

9.5 Future Research Opportunities

My thesis demonstrates the difficulty for researchers in examining associations between audit quality and market measures. While the problems with measures of audit quality are well known, I identify additional difficulties that can arise from model specification. Future research is needed to identify the precise mechanisms by which audit quality affects the predictability of earnings; however, in the absence of rigorous theoretical specifications, my study highlights the need to consider carefully the time horizons over which audit quality (or other variables posited to affect earnings quality) and the market proxy of interest are measured.

Second, I conjecture that the quality of analysts following a given firm is not random, and to my knowledge, the literature is relatively scarce on the determinants of this aspect of analyst following. Identifying the factors affecting a high-quality analyst's decision to follow a firm is important to understand the economic role of analyst quality on improvements in forecast accuracy. This is because previously documented evidence on analyst quality increasing forecast accuracy may be attributable to a high-quality analyst's superior ability to identify and cover a firm whose earnings can be easily predicted.

Further, my thesis shows that the relationship between auditor industry specialisation and forecast accuracy is more pronounced when client firm's operating risk is higher and analyst quality is lower. Future research may investigate other potential sources of predictable variation in the importance of audit quality in improving forecast accuracy.

Finally, I developed firm-year-level analyst quality metrics, which may be useful in testing the relative importance of other earnings quality proxies. In particular, future research may apply the difference in forecast accuracy between the 'worst' and 'best' quality analysts following a given firm to assess the effect of real and accrual-based earnings manipulation and disclosure quality.

REFERENCES

Abarbanell, J., and R. Lehavy. 2003. Biased forecasts or biased earnings? The role of reported earnings in explaining apparent bias and over/underreaction in analysts' earnings forecasts. *Journal of Accounting and Economics* 36 (1–3): 105–146.

Abbott, L. J., S. Parker, and G. Peters. 2005. *Auditor Industry Specialisation and Auditor Reporting*. Working paper, University of Memphis.

Affleck-Graves, J., C. M. Callahan, and N. Chipalkatti. 2002. Earnings predictability, information asymmetry, and market liquidity. *Journal of Accounting Research* 40 (3): 561–583.

Ahmed, A. S., S. M. K. Nainar, and J. Zhou. 2005. Do analysts' earnings forecasts fully reflect the information in accruals? *Canadian Journal of Administrative Sciences* 22 (4): 329–342.

Allayannis, G., B. Rountree, and J. P. Weston. 2005. *Earnings Volatility, Cash Flow Volatility, and Firm Value*. Working paper, University of Virginia.

Allen, E. J., C. R. Larson, and R. G. Sloan. 2013. Accrual reversals, earnings and stock returns. *Journal of Accounting and Economics* 56 (1): 113–129.

American Institute of Certified Public Accountants, Auditing Standards Board (AICPA). 1972. *Statement on Auditing Standards No. 1 Responsibilities and Functions of the Independent Auditor*. New York: AICPA.

Antle, R., and B. Nalebuff. 1991. Conservatism and auditor-client negotiations. *Journal of Accounting Research* 29: 31–54.

Arens, A. A., P. Best, G. Shailer, B. Fiedler, R. J. Elder, and M. Beasley. 2013. *Auditing, Assurance Services & Ethics in Australia: An Integrated Approach*. 9th ed. Frenchs Forest, NSW: Pearson Australia.

Ashbaugh, H., R. LaFond, and B. W. Mayhew. 2003. Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review* 78 (3): 611–639.

Ashton, A. H. 1991. Experience and error frequency knowledge as potential determinants of audit expertise. *The Accounting Review* 66 (2): 218–218.

Ashton, R. H., and J. J. Willingham. 1988. *Using and Evaluating Audit Decision Aids*. Working paper, University of Kansas.

Atiase, R. K. 1985. Predisclosure information, firm capitalization, and security price behaviour around earnings announcements. *Journal of Accounting Research* 23 (1): 21–36.

Bagnoli, M., S. G. Watts, and Y. Zhang. 2005. *Reg-FD and the Competitiveness of All-Star Analysts*. Working paper, Krannert Graduate School of Management.

Ball, R. 2013. *Accounting Informs Investors and Earnings Management is Rife: Two Questionable Beliefs*. Working paper, University of Chicago.

Ball, R., S. Jayaraman, and L. Shivakumar. 2012. Audited financial reporting and voluntary disclosure as complements: A test of the Confirmation Hypothesis. *Journal of Accounting and Economics* 53 (1–2): 136–166.

Balsam, S., J. Krishnan, and J. S. Yang. 2003. Auditor industry specialisation and earnings quality. *Auditing* 22 (2): 71–97.

Bandyopadhyay, S. P., and J. L. Kao. 2001. Competition and Big 6 brand name reputation: Evidence from the Ontario Municipal Audit Market. *Contemporary Accounting Research* 18 (1): 27–64.

Bannister, J. W., and H. A. Newman. 1996. Accrual usage to manage earnings toward financial analysts' forecasts. *Review of Quantitative Finance and Accounting* 7 (3): 259–278.

Barron, O. E., O. Kim, S. C. Lim, and D. E. Stevens. 1998. Using analysts' forecasts to measure properties of analysts' information environment. *The Accounting Review* 73 (4): 421–433.

Barth, M. E., D. P. Cram, and K. K. Nelson. 2001. Accruals and the prediction of future cash flows. *The Accounting Review* 76 (1): 27–58.

Basu, S. 1997. The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics* 24 (1): 3–37.

Basu, S., L.-S. Hwang, and C.-L. Jan. 2001. *Differences in Conservatism Between Big Eight and Non-Big Eight Auditors.* Working paper, Baruch College.

Baum, C. 2006. *An Introduction to Modern Econometrics using Stata*. College Station, TX: Stata Press.

Baum, C., M. E. Schaffer, and S. Stillman. 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. *The Stata Journal* 7 (4): 465–506.

Bauwhede, H. V., M. Willekens, and A. Gaeremynck. 2000. *Audit Quality, Public Ownership and Firms' Discretionary Accruals Management.* Working paper, Catholic University of Leuven.

Becker, C. L., M. L. Defond, J. Jiambalvo, and K. R. Subramanyam. 1998. The effect of audit quality on earnings management. *Contemporary Accounting Research* 15 (1): 1–24.

Bedard, J. C., and S. F. Biggs. 1991. Pattern recognition, hypotheses generation, and auditor performance in an analytical task. *The Accounting Review* 66 (3): 622–642.

Bedard, J. C., and M. T. H. Chi. 1993. Expertise in auditing. *Auditing: A Journal of Practice and Theory* 12: 21–21.

Bedard, J. C., and T. Mock. 1992. Expert and novice problem-solving behaviour in audit planning. *Auditing: A Journal of Practice and Theory* (Supplement): 1–20.

Behn, B., J. Choi, and T. Kang. 2008. Audit quality and properties of analyst earnings forecasts. *The Accounting Review* 83 (2): 327.

Benston, G. J. 1985. The market for public accounting services: Demand, supply and regulation. *Journal of Accounting and Public Policy* 4 (1): 33–79.

Bhattacharya, S. 2011. *Auditor Industry Specialisation and Earnings Response Coefficient: A New Zealand Perspective.* PhD Dissertation, Auckland University of Technology.

Bonner, S. E. 2008. *Judgment and Decision Making in Accounting*. New Jersey: Pearson Education.

Bonner, S. E., and B. L. Lewis. 1990. Determinants of auditor expertise. *Journal of Accounting Research* 28: 1–20.

Boone, J. P., I. K. Khurana, and K. K. Raman. 2012. Audit market concentration and auditor tolerance for earnings management. *Contemporary Accounting Research* 29 (4): 1171–1203.

Bradshaw, M. T., S. A. Richardson, and R. G. Sloan. 2001. Do analysts and auditors use information in accruals? *Journal of Accounting Research* 39 (1): 45–74.

Bradshaw, M. T., and R. G. Sloan. 2002. GAAP versus The Street: An empirical assessment of two alternative definitions of earnings. *Journal of Accounting Research* 40 (1): 41–66.

Brown, C. 2014. *The Lure of the Slant: Strategic Optimism and Asset Prices.* Working paper, National University of Singapore.

Brown, H. L., and A. M. Wright. 2008. Negotiation research in auditing. *Accounting Horizons* 22 (1): 91–109.

Brown, L. D., and E. Mohammad. 2010. Is analyst earnings forecast ability only firm specific?. *Contemporary Accounting Research* 27 (3): 727–750.

Burgstahler, D., and M. Eames. 2006. Management of earnings and analysts' forecasts to achieve zero and small positive earnings surprises. *Journal of Business Finance & Accounting* 33: 633–652.

Cambridge Advanced Learner's Dictionary. 2008. 3rd ed. Cambridge: Cambridge University Press.

Caramanis, C., and C. Lennox. 2008. Audit effort and earnings management. *Journal of Accounting and Economics* 45 (1): 116–138.

Carey, P., and R. Simnett. 2006. Audit partner tenure and audit quality. *The Accounting Review* 81 (3): 653.

Casey, R. J. 2012. *Do Independent Research Analysts Issue More Informative Recommendation Revisions?* Working paper, University of Illinois at Chicago.

Chan, L. K. C., D. Ikenberry, J. Lakonishok, and S. Lee. 2004. *Are All Analysts Equal? Consistency in Forecasting Ability.* Working paper, University of Illinois, Champaign.

Chang, C. J., and N-C. R. Hwang. 2003. The impact of retention incentives and client business risks on auditors' decisions involving aggressive reporting practices. *Auditing: A Journal of Practice & Theory* 22 (2): 207–218.

Charness, N., and R. S. Schultetus. 1999. Knowledge and expertise. In *Handbook of Applied Cognition*, edited by F. T. Durso, R. S. Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay & M. T. H. Chi, 57–81. Chichester, England: John Wiley & Sons Ltd.

Chase, W. G., and H. A. Simon. 1973. Perception in chess. *Cognitive Psychology* 4: 55–81.

Chen, C.-Y., C.-J. Lin, and Y.-C. Lin. 2008. Audit partner tenure, audit firm tenure, and discretionary accruals: Does long auditor tenure impair earnings quality? *Contemporary Accounting Research* 25 (2): 415–445.

Chen, K. Y., R. J. Elder, and J.-L. Liu. 2005. Auditor independence, audit quality and auditor-client negotiation outcomes: Some evidence from Taiwan. *Journal of Contemporary Accounting & Economics* 1 (2): 119–146.

Chen, S., S. Y. J. Sun, and D. Wu. 2010. Client importance, institutional improvements, and audit quality in China: An office and individual auditor level analysis. *The Accounting Review* 85 (1): 127–158.

Chi, M. T. H., R. Glaser, and E. Rees. 1982. Expertise in problem solving. In *Advances in the Psychology of Human Intelligence*, edited by R. J. Sternberg, 7–75. vol 1. Hillsdale, NJ: Erlbaum.

Chin, C.-L., and H.-Y. Chi. 2009. Reducing restatements with increased industry expertise. *Contemporary Accounting Research* 26 (3): 729–765.

Choi, J.-H., and Y. K. Kwon. 2008. Theory on the association between audit quality and the accuracy and dispersion of analysts' earnings forecasts. *Seoul Journal of Business* 14 (2): 93–107.

Clarkson, P. M. 2000. Auditor quality and the accuracy of management earnings forecasts. *Contemporary Accounting Research* 17 (4): 595–622.

Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics* 27 (3): 285–303.
Clement, M. B., L. Koonce, and T. J. Lopez. 2007. The roles of task-specific forecasting experience and innate ability in understanding analyst forecasting performance. *Journal of Accounting and Economics* 44 (3): 378–398.
Clement, M. B., L. Rees, and E. P. Swanson. 2003. The influence of culture and corporate governance on the characteristics that distinguish superior analysts. *Journal of Accounting, Auditing & Finance* 18 (4): 593–618.
Clement, M. B., and S. Y. Tse. 2003. Do investors respond to analysts' forecast revisions as if forecast accuracy is all that matters? *The Accounting Review* 78 (1): 227–249.
Clement, M. B., and S. Y. Tse. 2005. Financial analyst characteristics and herding behaviour in forecasting. *The Journal of Finance* 60 (1): 307–341.
Collins, D. W., S. P. Kothari, and J. D. Rayburn. 1987. Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics* 9 (2): 111–138.
Crabtree, A. D., and J. J. Maher. 2005. Earnings predictability, bond ratings, and bond yields. *Review of Quantitative Finance and Accounting* 25 (3): 233–253.
Cram, D. P., V. Karan, and I. Stuart. 2009. Three threats to validity of choice-based and matched-sample studies in accounting research. *Contemporary Accounting Research* 26 (2): 477–516.
Craswell, A. T., J. R. Francis, and S. L. Taylor. 1995. Auditor brand name reputations and industry specialisations. *Journal of Accounting and Economics* 20 (3): 297–322.
Das, S., C. B. Levine, and K. Sivaramakrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73 (2): 277–294.
Davidson, R. A., and D. Neu. 1993. A note of the association between audit firm size and audit. *Contemporary Accounting Research* 9 (2): 479.
DeAngelo, L. E. 1981. Auditor size and audit quality. *Journal of Accounting and Economics* 3 (3): 183–199.
Dechow, P. M., and I. Dichev, D. 2002. The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77: 35–59.
Dechow, P. M., S. P. Kothari, and R. L. Watts. 1998. The relation between earnings and cash flows. *Journal of Accounting and Economics* 25 (2): 133–168.
Dechow, P. M., S. A. Richardson, and I. Tuna. 2003. Why are earnings kinky? An examination of the earnings management explanation. *Review of Accounting Studies* 8 (2): 355–384.
DeFond, M. L., J. R. Francis, and T. J. Wong. 2000. Auditor industry specialisation and market segmentation: Evidence from Hong Kong. *Auditing: A Journal of Practice & Theory* 19 (1): 49–66.
DeFond, M. L., K. Raghunandan, and K. R. Subramanyam. 2002. Do non-audit service fees impair auditor independence? Evidence from going concern audit opinions. *Journal of Accounting Research* 40 (4): 1247–1274.
DeFond, M. L., and J. Zhang. 2014. *A Review of Archival Auditing Research.* Working paper, University of Southern California.
Dhaliwal, D. S., S. Radhakrishnan, A. Tsang, and Y. G. Yang. 2012. Nonfinancial disclosure and analyst forecast accuracy: International evidence on corporate social responsibility disclosure. *The Accounting Review* 87 (3): 723–759.
Dichev, I. D., and V. W. Tang. 2009. Earnings volatility and earnings predictability. *Journal of Accounting and Economics* 47 (1–2): 160–181.

Dowling, C., and S. Leech. 2007. Audit support systems and decision aids: Current practice and opportunities for future research. *International Journal of Accounting Information Systems* 8 (2): 92–116.

Drake, M., and L. Myers. 2011. Analysts' accrual-related over-optimism: Do analyst characteristics play a role? *Review of Accounting Studies* 16 (1): 59–88.

Dunn, K., and B. Mayhew. 2004. Audit firm industry specialisation and client disclosure quality. *Review of Accounting Studies* 9 (1): 35–58.

Eames, M. J., and S. M. Glover. 2003. Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review* 78 (3): 707–724.

Eleswarapu, V. R., R. Thompson, and K. Venkataraman. 2004. The impact of regulation fair disclosure: Trading costs and information asymmetry. *The Journal of Financial and Quantitative Analysis* 39 (2): 209–225.

Emery, D. R., and X. Li. 2009. Are the Wall Street analyst rankings popularity contests? *Journal of Financial and Quantitative Analysis* 44 (2): 411–437.

Ernst&Young (EY). 2014. Productivity: It's on Our Agenda. Is It on Yours? Available at: http://www.ey.com/AU/en/Services/Advisory.

European Commission. 2010. Green Paper. Audit Policy: Lessons from the Crisis. COM(2010) 561 Final. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0561:FIN:EN:PDF

Ferguson, A., J. R. Francis, and D. J. Stokes. 2003. The effects of firm-wide and office-level industry expertise on audit pricing. *The Accounting Review* 78 (2): 429–448.

Financial Accounting Standards Board. (2010). Conceptual framework for financial reporting. *Statement of Financial Accounting Concepts No. 8.*

Financial Reporting Council (FRC). 2006. Promoting Audit Quality. Discussion paper. Available at: http://www.frc.org.uk/Our-Work/Publications/FRC-Board/Discussion-Paper-Promoting-Audit-Quality.aspx.

Financial Reporting council (FRC). 2008. The Audit Quality Framework. Available at: http://www.frc.org.uk/Our-Work/Publications/FRC-Board/The-Audit-Quality-Framework-%281%29.aspx.

Findlay, S., and P. G. Mathew. 2006. An examination of the differential impact of regulation FD on analysts' forecast accuracy. *Financial Review* 41 (1): 9–31.

Fombrun, C. J., N. A. Gardberg, and J. M. Sever. 2000. The reputation quotient: A multi-stakeholder measure of corporate reputation. *The Journal of Brand Management* 7 (4): 241–255.

Francis, J. R. 2011. A framework for understanding and researching audit quality. *Auditing: A Journal of Practice & Theory* 30 (2): 125–152.

Francis, J. R., R. LaFond, P. Olsson, and K. Schipper. 2005a. The market pricing of accruals quality. *Journal of Accounting and Economics* 39: 295–327.

Francis, J. R., E. L. Maydew, and H. C. Sparks. 1999. The role of Big 6 auditors in the credible reporting of accruals. *Auditing: A Journal of Practice & Theory* 18 (2): 18–34.

Francis, J. R., P. N. Michas, and S. E. Seavey. 2012. Does audit market concentration harm the quality of audited earnings? Evidence from audit markets in 42 countries. *Contemporary Accounting Research* 30 (1): 325–355.

Francis, J. R., K. Reichelt, and D. Wang. 2005b. The pricing of national and city-specific reputations for industry expertise in the U.S. audit market. *The Accounting Review* 80 (1): 113–136.

Francis, J. R., D. J. Stokes, and D. Anderson. 1999. City markets as a unit of analysis in audit research and the re-examination of Big 6 market shares. *Abacus* 35 (2): 185–206.

Francis, J. R., and E. R. Wilson. 1988. Auditor changes: A joint test of theories relating to agency costs and auditor differentiation. *The Accounting Review* 63 (4): 663–682.

Francis, J. R., and M. D. Yu. 2009. Big 4 office size and audit quality. *The Accounting Review* 84 (5): 1521–1552.

Frankel, R. M., M. F. Johnson, and K. K. Nelson. 2002. The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review* 77 (Supplement):71–105.

Frederick, D. M., and R. Libby. 1986. Expertise and auditors' judgments of conjunctive events. *Journal of Accounting Research* 24 (2): 270–290.

Gaver, J. J., and J. S. Paterson. 2007. The influence of large clients on office-level auditor oversight: Evidence from the property-casualty insurance industry. *Journal of Accounting and Economics* 43 (2–3): 299–320.

General Accounting Office. 2003a. Public Accounting Firms: Mandated Study on Consolidation and Competition. Available at: http://www.gao.gov/new.items/d03864.pdf.

General Accounting Office. 2003b. Public Accounting Firms: Required Study on the Potential Effects of Mandatory Audit Firm Rotation. Available at: http://www.gao.gov/products/GAO-04-216.

General Accounting Office. 2008. Audits of Public Companies Continued Concentration in Audit Market for Large Public Companies does not Call for Immediate Action. Available at: http://www.gao.gov/products/GAO-08-163.

Gibbins, M., S. McCracken, and S. Salterio. 2003. *Auditor-Client Management Negotiation Concerning Client's Financial Reporting: Evidence From the Client's Side.* Working paper, University of Alberta.

Gibbins, M., S. Salterio, and A. Webb. 2001. Evidence about auditor-client management negotiation concerning client's financial reporting. *Journal of Accounting Research* 39 (3): 535–563.

Gilson, S. C., P. M. Healy, C. F. Noe, and K. G. Palepu. 2001. Analyst specialisation and conglomerate stock breakups. *Journal of Accounting Research* 39 (3): 565–582.

Givoly, D., and C. Hayn. 2000. The changing time-series properties of earnings, cash flows and accruals: Has financial reporting become more conservative? *Journal of Accounting and Economics* 29 (3): 287–320.

Godfrey, J. M., and J. Hamilton. 2005. The impact of R&D intensity on demand for specialist auditor services. *Contemporary Accounting Research* 22 (1): 55–93.

Gramling, A. A., J. Krishnan, and Y. Zhang. 2011. Are PCAOB-identified audit deficiencies associated with a change in reporting decisions of triennially inspected audit firms? *Auditing: A Journal of Practice & Theory* 30 (3): 59–79.

Gramling, A. A., and D. N. Stone. 2001. Audit firm industry expertise: A review and synthesis of the archival literature. *Journal of Accounting Literature* 20: 1–29.

Greene, W. H. 2002. *Econometric Analysis.* Upper Saddle River, New Jersey: Prentice Hall.

Gu, Z., and J. S. Wu. 2003. Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics* 35 (1): 5–29.

Gul, F. A., S. Y. K. Fung, and B. Jaggi. 2009. Earnings quality: Some evidence on the role of auditor tenure and auditors' industry expertise. *Journal of Accounting and Economics* 47 (3): 265–287.

Guo, S., and M. Fraser. 2010. *Propensity Score Analysis: Statistical Methods and Applications*. Thousand Oaks, CA: SAGE Publications, Inc.

Habib, A., and M. B. U. Bhuiyan. 2011. Audit firm industry specialisation and the audit report lag. *Journal of International Accounting, Auditing and Taxation* 20 (1): 32–44.

Hall, J. L., and P. B. Tacon. 2010. Forecast accuracy and stock recommendations. *Journal of Contemporary Accounting & Economics* 6 (1): 18–33.

Hamilton, R. E., and W. F. Wright. 1982. Internal control judgments and effects of experience: Replications and extensions. *Journal of Accounting Research* 20 (2): 756–765.

Hayes-Roth, F., and D. B. Lenat. 1983. What are expert systems? In *Building Expert Systems,* edited by F. Hayes-Roth, D. A. Waterman, D. B. Lenat, 31–59. California: Addison-Wesley.

He, W., B. Sidhu, and S. Taylor. 2011. *Audit Quality and Properties of Analysts' Information Environment.* Working paper, University of New South Wales.

Healy, P. M., and J. M. Wahlen. 1999. A review of the earnings management literature and its implications for standard setting. *Accounting Horizons* 13 (4): 365–383.

Heflin, F., K. R. Subramanyam, and Y. Zhang. 2003. Regulation FD and the financial information environment: Early evidence. *The Accounting Review* 78 (1): 1–37.

Hogan, C. E., and D. C. Jeter. 1999. Industry specialisation by auditors. *Auditing: A Journal of Practice & Theory* 18 (1).

Hope, O.-K. 2003. Disclosure practices, enforcement of accounting standards, and analysts' forecast accuracy: An international study. *Journal of Accounting Research* 41 (2): 235–272.

House of Lords. 2010a. Call for Evidence: Auditors: Market Concentration and their Role. Available at: http://www.parliament.uk/documents/lords-committees/economic-affairs/auditors/cfeauditors20100727.pdf.

House of Lords. 2010b. Report (volume I) Auditors: Market Concentration and their Role. Available at: http://www.publications.parliament.uk/pa/ld201011/ldselect/ldeconaf/119/119.pdf.

House of Lords. 2010c. Report (Volume II) Auditors: Market Concentration and their Role. Available at: http://www.publications.parliament.uk/pa/ld201011/ldselect/ldeconaf/119/119ii.pdf.

Houston, R. W. 1999. The effects of fee pressure and client risk on audit seniors' time budget decisions. *Auditing: A Journal of Practice & Theory* 18 (2): 70–86.

Hwang, L., C. Jan, and S. Basu. 1996. Loss firms and analysts' earnings forecast errors. *The Journal of Financial Statement Analysis* 1 (Winter): 18–29.

Imhoff, E. A., Jr., and G. J. Lobo. 1992. The effect of ex ante earnings uncertainty on earnings response coefficients. *The Accounting Review* 67 (2): 427–439.

Institutional Investor (*II*). October 2010. Methodology. Available at: http://www.institutionalinvestor.com/Research/4569/Methodology.html.

International Organization of Securities Commissions (IOSCO). 2009. Consultation Report on Transparency of Firms that Audit Public Companies. Available at: http://www.iosco.org/library/pubdocs/pdf/IOSCOPD302.pdf.

Jacob, J., T. Z. Lys, and M. A. Neale. 1999. Expertise in forecasting performance of security analysts. *Journal of Accounting and Economics* 28 (1): 51–82.

Janvrin, D., J. Bierstaker, and D. J. Lowe. 2008. An examination of audit information technology use and perceived importance. *Accounting Horizons* 22 (1): 1–21.

Jarmon, R. 2009. Reputation's effect on pricing power: The importance of strategy. *Corporate Reputation Review* 12 (3): 281–296.

Jayaraman, S. 2008. Earnings volatility, cash flow volatility, and informed trading. *Journal of Accounting Research* 46 (4): 809–851.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the firm: Managerial behaviour, agency costs and ownership structure. *Journal of Financial Economics* 3 (4): 305–360.

Jones, J. J. 1991. Earnings management during import relief investigations. *Journal of Accounting Research* 29 (2): 193–228.

Kennedy, P. 1992. *A Guide to Econometrics*. Oxford: Blackwell.

Keskek, S., L. A. Myers, T. C. Omer, and M. Shelley. 2013. *A Re-examination of the Importance of Analyst and Forecast Characteristics for Forecast Accuracy: The Effects of Analyst Disagreement, Stale Forecasts, and the Information Environment.* Working paper, University of Arkansas.

Khurana, I. K., and K. K. Raman. 2004. Litigation risk and the financial reporting credibility of Big 4 versus Non-Big 4 audits: Evidence from Anglo-American countries. *The Accounting Review* 79 (2): 473–495.

Kim, Y., G. J. Lobo, and M. Song. 2011. Analyst characteristics, timing of forecast revisions, and analyst forecasting ability. *Journal of Banking & Finance* 35 (8): 2158–2168.

Kini, O., S. Mian, M. Rebello, and A. Venkateswaran. 2009. On the structure of analyst research portfolios and forecast accuracy. *Journal of Accounting Research* 47 (4): 867–909.

Kleibergen, F. 2007. Generalizing weak instrument robust IV statistics towards multiple parameters, unrestricted covariance matrices and identification statistics. *Journal of Econometrics* 139 (1): 181–216.

Kleibergen, F., and R. Paap. 2006. Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* 133 (1): 97–126.

Knechel, W. R., G. V. Krishnan, M. Pevzner, L. B. Shefchik, and U. K. Velury. 2013. Audit quality: Insights from the academic literature. *Auditing: A Journal of Practice & Theory* 32 (Supplement 1): 385–421.

Knechel, W. R., V. Naiker, and G. Pacheco. 2007. Does auditor industry specialisation matter? Evidence from market reaction to auditor switches. *Auditing: A Journal of Practice & Theory* 26 (1): 19–45.

Kothari, S. P. 2001. Capital markets research in accounting. *Journal of Accounting and Economics* 31 (1–3): 105–231.

KPMG. 2014. Industries We Operate In. Available at: http://www.kpmg.com/global/en/careers/whatcanido/industries/Pages/default.aspx.

Krishnan, G. V. 2003. Audit quality and the pricing of discretionary accruals. *Auditing: A Journal of Practice & Theory* 22 (1): 109–126.

Krishnan, G. V. 2005. The association between Big 6 auditor industry expertise and the asymmetric timeliness of earnings. *Journal of Accounting, Auditing & Finance* 20 (3): 209–228.

Krishnan, J. 2001. A comparison of auditors' self-reported industry expertise and alternative measures of industry specialisation. *Asia-Pacific Journal of Accounting & Economics* 8 (2): 127–142.

Kross, W., B. Ro, and D. Schroeder. 1990. Earnings expectations: The analysts' information advantage. *The Accounting Review* 65 (2): 461–476.

Kwon, S. 1996. The impact of competition within the client's industry on the auditor selection decision. *Auditing: A Journal of Practice & Theory* 15 (1): 53–69.

Lang, M. H., and R. J. Lundholm. 1996. Corporate disclosure policy and analyst behaviour. *The Accounting Review* 71 (4): 467–492.

Larcker, D. F., and S. A. Richardson. 2004. Fees paid to audit firms, accrual choices, and corporate governance. *Journal of Accounting Research* 42 (3): 625–658.

Larcker, D. F., and S. A. Richardson., and T. O. Rusticus. 2010. On the use of instrumental variables in accounting research. *Journal of Accounting and Economics* 49 (3): 186–205.

Lawrence, A., M. Minutti-Meza, and P. Zhang. 2011. Can Big 4 versus Non-Big 4 differences in audit-quality proxies be attributed to client characteristics? *The Accounting Review* 86 (1): 259–286.

Lee, S. 2004. *Consistency in Analyst Forecasting Ability.* Working paper, University of Illinois at Urbana-Champaign.

Lennox, C. S., J. R. Francis, and Z. Wang. 2012. Selection models in accounting research. *The Accounting Review* 87 (2): 589–616.

Leuz, C., D. Nanda, and P. D. Wysocki. 2003. Earnings management and investor protection: An international comparison. *Journal of Financial Economics* 69 (3): 505–527.

Levitt, A. 1998. The importance of high quality accounting standards. *Accounting Horizons* 12 (2): 79–82.

Lim, C.-Y., and H.-T. Tan. 2008. Non-audit service fees and audit quality: The impact of auditor specialisation. *Journal of Accounting Research* 46 (1): 199–246.

Loh, R. K., and R. M. Stulz. 2010. When are analyst recommendation changes influential? *Review of Financial Studies* 24 (2): 593-627.

Low, K.-Y. 2004. The effects of industry specialisation on audit risk assessments and audit-planning decisions. *The Accounting Review* 79 (1): 201–219.

Lys, T., and S. Sohn. 1990. The association between revisions of financial analysts' earnings forecasts and security-price changes. *Journal of Accounting and Economics* 13 (4): 341–363.

Marcus, K., D. A. Reppenhagen, and J. W. Tucker. 2014. Meeting individual analyst expectations. *The Accounting Review* Forthcoming.

Marquardt, D. W. 1970. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* 12:256–591.

McNichols, M. F. 2002. Discussion of the quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review* 77 (Supplement): 61–69.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1997. Do security analysts improve their performance with experience? *Journal of Accounting Research* 35: 131–157.

Minton, B. A., C. M. Schrand, and B. R. Walther. 2002. The role of volatility in forecasting. *Review of Accounting Studies* 7: 195–215.

Minutti-Meza, M. 2013. Does auditor industry specialisation improve audit quality? *Journal of Accounting Research* 51 (4): 779–817.

Moroney, R. 2007. Does industry expertise improve the efficiency of audit judgment? *Auditing: A Journal of Practice & Theory* 26 (2): 69–94.

Moroney, R., and P. Carey. 2011. Industry- versus task-based experience and auditor performance. *Auditing: A Journal of Practice & Theory* 30 (2): 1–18.

Neal, T. L., and R. R. Riley. 2004. Auditor industry specialist research design. *Auditing: A Journal of Practice & Theory* 23 (2): 169–177.

Neter, J., W. Wasserman, and M. H. Kutner. 1989. *Applied Linear Regression Models.* Homewood, IL: Irwin.

Numan, W., and M. Willekens. 2012. *Competitive Pressure, Audit Quality and Industry Specialisation.* Working paper, KU Leuven.

O'Brien, R. M. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity* 41: 673–690.

O'Donnell, E., and J. J. J. Schultz. 2003. The influence of business-process-focused audit support software on analytical procedures judgments. *Auditing: A Journal of Practice & Theory* 22 (2): 265–279.

Ohlson, J. A. 1990. A synthesis of security valuation theory and the role of dividends, cash flows, and earnings. *Contemporary Accounting Research* 6 (2): 648–676.

O'Keefe, T. B., D. A. Simunic, and M. T. Stein. 1994. The production of audit services: Evidence from a major public accounting firm. *Journal of Accounting Research* 32 (2): 241–261.

Oskamp, S. 1965. Overconfidence in case-study judgements. *The Journal of Consulting Psychology* 2: 261–265.

Oxford Dictionary of English. 2010. 3rd ed. Oxford, England: Oxford University Press.

Palmrose, Z.-V. 1988. Competitive manuscript co-winner: An analysis of auditor litigation and audit service quality. *The Accounting Review* 63 (1): 55–73.

Payne, J. 2008. The influence of audit firm specialisation on analysts' forecast errors. *Auditing: A Journal of Practice & Theory* 27 (2): 109.

Peters, G., L. J. Abbott, and S. Parker. 2001. *Voluntary Disclosures and Auditor Specialisation: The Case of Commodity Derivative Disclosures.* Working paper, University of Georgia.

Petersen, M. A. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies* 22 (1): 435–480.

PricewaterhouseCoopers (PWC). 2014. Our Commitment: Deliver Value, with You, Every Day. Available at: http://www.pwc.com/us/en/about-us/index.jhtml.

Reichelt, K. J., and D. Wang. 2010. National and office-specific measures of auditor industry expertise and effects on audit quality. *Journal of Accounting Research* 48 (3): 647–686.

Reynolds, J. K., and J. R. Francis. 2001. Does size matter? The influence of large clients on office-level auditor reporting decisions. *Journal of Accounting and Economics* 30 (3): 375–400.

Romanus, R. N., J. J. Maher, and D. M. Fleming. 2008. Auditor industry specialisation, auditor changes, and accounting restatements. *Accounting Horizons* 22 (4): 389–413.

Rosenbaum, P. R., and D. B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70 (1): 41–55.

Sartori, A. E. 2003. An estimator for some binary-outcome selection models without exclusion restrictions. *Political Analysis* 11 (2): 111–138.

Schelleman, C., and W. R. Knechel. 2010. Short-term accruals and the pricing and production of audit services. *Auditing: A Journal of Practice & Theory* 29 (1): 221–250.

Schipper, K., and L. Vincent. 2003. Earnings quality. *Accounting Horizons* 17: 97–110.

Siegel, P. H., J. P. Lessard, and K. E. Karim. 2011. Analyst forecast accuracy and firm growth. *Advances in Quantitative Analysis of Finance and Accounting* 9: 1–31.

Simon, D. P., and H. A. Simon. 1978. Individual differences in solving physics problems. In *Children's Thinking: What Develops?*, edited by R. Siegkr, 325–348. Hillsdale, NJ: Erlbaum.

Simunic, D. A., and M. T. Stein. 1987. *Product Differentiation in Auditing: Auditor Choice in the Market for Unseasoned New Issues*. Vancouver, B.C: Canadian Certified General Accountants' Research Foundation.

Skinner, D., and R. Sloan. 2002. Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio. *Review of Accounting Studies* 7 (2): 289–312.

Sloan, R. G. 1996. Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review* 71 (3): 289–315.

Smith, A. 1776. *The Wealth of Nations*. London: W. Strahan and T. Cadell.

Solomon, I., M. D. Shields, and O. R. Whittington. 1999. What do industry-specialist auditors know? *Journal of Accounting Research* 37 (1): 191–208.

Stice, J. D. 1991. Using financial and market information to identify pre-engagement factors associated with lawsuits against auditors. *The Accounting Review* 66 (3): 516–533.

Stickel, S. E. 1992. Reputation and performance among security analysts. *The Journal of Finance* 47 (5): 1811–1836.

Stock, J. H., J. H. Wright, and M. Yogo. 2002. A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics* 20: 518–529.

Subramanyam, K. R. 1996. The pricing of discretionary accruals. *Journal of Accounting and Economics* 22 (1–3): 249–281.

Taylor, J. B., and L. Frost. 2009. *Microeconomics*. 4th ed. Milton, QLD: John Wiley & Sons Australia.

Taylor, M. H. 2000. The effects of industry specialisation on auditors' inherent risk assessments and confidence judgements. *Contemporary Accounting Research* 17 (4): 693–712.

Teoh, S. H., and T. J. Wong. 1993. Perceived auditor quality and the earnings response coefficient. *The Accounting Review* 68 (2): 346–366.

Thompson, S. B. 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99: 1–10.

Watkins, A. L., W. Hillison, and S. E. Morecroft. 2004. Audit quality: A synthesis of theory and empirical evidence *Journal of Accounting Literature* 23: 153–153.

Watts, R. L., and J. L. Zimmerman. 1978. Towards a positive theory of the determination of accounting standards. *The Accounting Review* 53 (1): 112–134.

Wharton Research Data Services (WRDS). 2010. Summary and Data Checks for the IBES Update (September 2010), Data Comparison: September 2010 vs. April 2010 Updates. Available at: https://wrds-web.wharton.upenn.edu/wrds/support/Data/_001Manuals%20and%20Overviews/_003I-B-E-S/_004Release%20Notes/_994Summary%20and%20Comparison%20of%20IBES%20Update%202010-09.cfm.

Wilks, T. J., and M. F. Zimbelman. 2004. Decomposition of fraud-risk assessments and auditors' sensitivity to fraud cues. *Contemporary Accounting Research* 21 (3): 719–745.

Wilson, M. K., and Y. Wu. 2011. Do publicly signalled earnings management incentives affect analyst forecast accuracy? *Abacus* 47 (3): 315–342.

Wooldridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. London: MIT Press.

Wooldridge, J. M. 2008. *Introductory Econometrics: A Modern Approach*. 4th ed. Mason, Ohio: South-Western Cengage Learnings.

Wright, A., and S. Wright. 1997. The effect of industry experience on hypothesis generation and audit planning decisions. *Behavioural Research in Accounting* 9: 273–294.

Xie, H. 2001. The mispricing of abnormal accruals. *The Accounting Review* 76 (3): 357–373.

Yardley, J. A., N. L. Kauffman, T. D. Cairney, and W. D. Albrecht. 1992. Supplier behaviour in the U.S. audit market. *Journal of Accounting Literature* 11: 151–151.

Zhang, X. F. 2006. Information uncertainty and stock returns. *The Journal of Finance* 61 (1): 105–137.

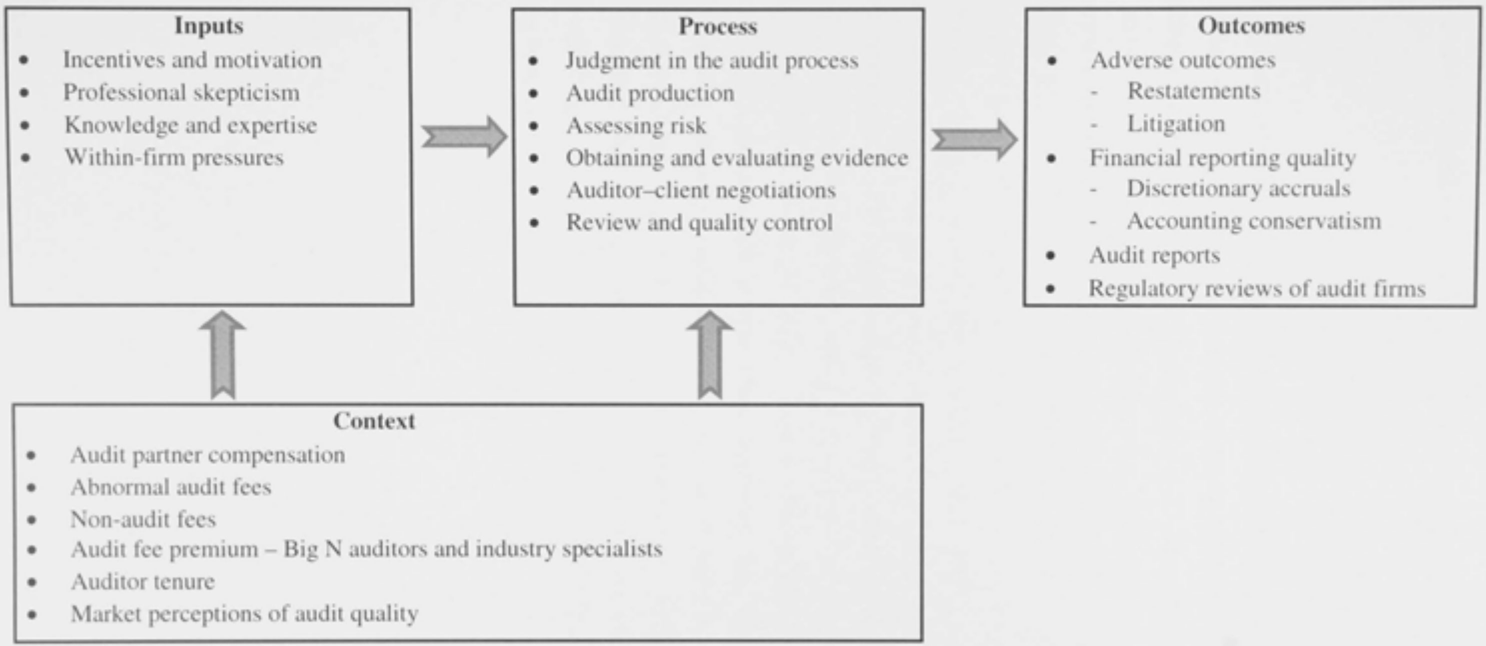
Zivot, E., R. Startz, and C. R. Nelson. 1998. Valid confidence intervals and inference in the presence of weak instruments. *International Economic Review* 39 (4): 1119–1144.

Zmijewski, M. E. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* 22: 59–82.

APPENDICES

Appendix A: KKPSV (2013) Audit Quality Framework

Figure A.1 KKPSV Audit Quality Framework



Appendix B: Probit Regressions of Benchmark-Beating against Audit Quality

Below I report the results of probit regressions modelling the likelihood of zero or small positive forecast errors to provide further evidence that long-horizon forecasts are relatively independent of earnings management behaviour. The dependent variable in the probit regression is a dummy variable (*MBE*) indicating that the signed forecast error is between 0 and 1 cent inclusive. The test variable is either a continuous or dummy measure of industry specialisation (*INDSP*). If firms audited by an industry specialist are constrained in their attempts to manipulate earnings to meet or beat analysts' forecasts, industry specialisation should be negatively correlated with the likelihood of meeting or beating analysts' forecasts. Table B.1 reports the probit regressions of *MBE* against audit quality and controls for both short- and long-horizon forecasts. Consistent with Payne's (2008) results, both continuous and dichotomous measures of industry specialisation (*INDSP*) are negatively associated with short-horizon defined *MBE*, suggesting that industry specialist audit firms may constrain benchmark-beating earnings management. If, however, *MBE* is measured as a function of long-horizon forecast errors, *INDSP* is not significantly correlated with any extant benchmark beating, and thus the relation between *INDSP* and long-horizon accuracy is less likely to be confounded by benchmark-beating behaviour.

Table B.1: Probit Regressions of the Likelihood of Meeting or Just Beating Analysts' Consensus Earnings Forecasts

Dependent Variable: *MBE*

Columns		(1)	(2)	(3)	(4)
		Short-Horizon Forecasts		Long-Horizon Forecasts	
	Pred.	*INDSP_cont*	*INDSP_dum*	*INDSP_cont*	*INDSP_dum*
INDSP	–	−0.9150**	−0.0891***	−0.4470	0.0086
		(0.026)	(<0.001)	(0.554)	(0.902)
DISP	–	−0.2210***	−0.2200***	−0.0752	−0.0746
		(<0.001)	(<0.001)	(0.213)	(0.216)
SIZE	?	−0.0899***	−0.0904***	−0.0390**	−0.0395***
		(<0.001)	(<0.001)	(0.011)	(0.010)
NUMEST	+	0.0214***	0.0215***	0.0101***	0.0100***
		(<0.001)	(<0.001)	(0.001)	(0.001)
ABSACCR	–	−0.5550***	−0.5550***	−0.6450**	−0.6460**
		(<0.001)	(<0.001)	(0.020)	(0.020)
LOSS	–	−0.2650***	−0.2620***	−0.0730	−0.0750
		(<0.001)	(<0.001)	(0.269)	(0.257)
ABSECHG	–	−2.3450***	−2.3460***	−2.6750***	−2.670***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
PERSIST	?	−0.0505***	−0.0502***	−0.0637*	−0.0637*
		(0.006)	(0.007)	(0.081)	(0.080)
YEAR		yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes
CONSTANT		−1.135***	−1.180***	−1.425***	−1.426***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
Cases where *MBE*=1		5,830	5,830	663	663
% of meet or beat		19.29%	19.29%	2.86%	2.86%
N		30,225	30,225	23,203	23,203
Pseudo R^2		0.063	0.063	0.043	0.043

Robust *p*-values of the coefficients are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results of probit regressions modelling the likelihood of zero or small positive short-horizon forecast errors (Columns 1 and 2) and long-horizon forecast errors (Columns 3 and 4).

Variable Definitions: *MBE* equals 1 if the signed analyst forecast error is between 0 and 1 cent inclusive, 0 otherwise; *INDSP_cont* is the continuous measure of portfolio-share audit firm industry specialisation, measured as the sum of the square root of the total assets of the clients that an audit firm services in a specific industry divided by the sum of the square root of the total assets of all clients of that audit firm; *INDSP_dum* is the dichotomous measure of portfolio-share audit firm industry specialisation, equals 1 if *INDSP_cont* > (3 / number of two-digit industry codes used in the analysis in any given year), 0 otherwise; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *SIZE* is the natural log of total assets; *NUMEST* is the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ABSACCR* is the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *PERSIST* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise; *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

Appendix C: Sample Selection for Tests of H1a (Short-Horizon Forecasts)

Table C.1: Sample Selection for Tests of H1a (Short-Horizon Forecasts)

	Payne-type models		BCK-type models	
Available **short-horizon** consensus forecasts		65,710		65,710
less: unavailable financial information from *COMPUSTAT*	−11,973	53,737	−24,358	41,352
less: non-Big N firms	−5,646	48,091	−3,439	37,913
less: financial sector firms	−5,217	42,874	−810	37,103
less: firms in industries with less than 20 members in a given year	−615	42,259	−545	36,558
less: firms subject to modified audit opinions	−15	42,244	−13	36,545
less: observations where forecast dispersion cannot be calculated	−9,814	32,430	−6,329	30,216
less: extreme observations for continuous variable (1/99 percentile)	−624		−815	
Final sample of short-horizon forecast (used in single-stage models)		<u>31,806</u>		<u>29,401</u>
less: additional data requirements for the endogeneity-corrected regressions	−448		−700	
Final sample of short-horizon forecast (used in 2SLS)		<u>31, 358</u>		<u>28,701</u>

Appendix D: First Stage Regressions (Logit Regressions) of Propensity Score Matching (PSM) Models

Table D.1: First Stage (Logit Regressions) of PSM Models

Dependent Variable: *INDSP_dum*

Columns	Pred.	(1) Test of H1b (BCK Model)	(2) Test of H1b (Payne Model)	(3) Test of H2	(4) Tests of H3a	(5) Tests of H3b (All-Star Proxy)	(6) Test of H3b (excl. STAR proxy)
DISP	?	0.4213	2.7632		−0.4839	0.8509	1.3702
		(0.816)	(0.122)		(0.815)	(0.877)	(0.646)
HORIZON	?	−0.1121		−0.2144	−0.2740		
		(0.598)		(0.366)	(0.253)		
SIZE	+	0.1371***	0.0574***	0.0741***	0.1156***	0.1922***	0.1284***
		(<0.001)	(0.001)	(0.001)	(<0.001)	(<0.001)	(<0.001)
NUMEST	?	0.0056	0.0111***	0.0375	−0.0568	−0.1316*	−0.2514**
		(0.885)	(0.003)	(0.368)	(0.196)	(0.072)	(0.010)
ZSCORE	?	−0.0269		−0.0059	−0.0312	−0.0669*	−0.0264
		(0.154)		(0.761)	(0.141)	(0.056)	(0.251)
LOSS	−	0.2447***	0.2818***	0.2274**	0.1714*	0.1234	0.1144
		(0.005)	(<0.001)	(0.011)	(0.083)	(0.483)	(0.279)
ABSECHG	−	−1.3078***	−0.5972		−0.8768*	0.1090	−0.7746
		(0.003)	(0.173)		(0.087)	(0.916)	(0.196)
STDROE	+	2.1418***			1.9842***	2.7877***	2.1559***
		(<0.001)			(<0.001)	(<0.001)	(<0.001)
EL	−	−0.0373		−0.0093	0.0036	0.0141	−0.0184
		(0.170)		(0.751)	(0.908)	(0.743)	(0.567)
ABSACCR	−		−0.3820				
			(0.212)				
PERSIST	−		0.0328				
			(0.500)				
FFOLLOW	?				0.0228***		
					(<0.001)		
IFOLLOW	?				−0.0485***		
					(<0.001)		
DIFANQ	?						−0.0052

							(0.696)
ANQ_B	?						0.0125
							(0.293)
HORIZON_B	?					−0.0018	−0.0007
						(0.332)	(0.304)
DIFHORIZON	?					−0.0034	−0.0009
						(0.141)	(0.324)
YEAR		yes	yes	yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes	yes	yes
CONSTANT		1.3635	1.2478***	2.7486**	2.7927**	1.1810*	1.5369***
		(0.269)	(<0.001)	(0.045)	(0.046)	(0.073)	(<0.001)
N		19,599	20,765	16,864	16,720	5,565	15,098
Pseudo R^2		0.5465	0.534	0.5538	0.5815	0.4818	0.5814
ROC Area		0.9065	0.9044	0.907	0.9096	0.9105	0.9095

P-values of the coefficients are two-tailed reported parentheses (*** p<0.01, ** p<0.05, * p<0.1).

This table presents the results for the first-stage regressions of the PSM models. Columns 1 and 2 report the regressions for tests of H1b based on the BCK-type and Payne-type models. Column 3 reports the regression for tests of H2. Column 4 reports the regression for tests of H3a. Column 5 reports the regression for tests of H3b based on the model using the 'All-Star' analyst quality proxy to identify the 'best' and 'worst' analysts. Column 6 reports the regression for tests of H3b based on all the other models.

Variable Definition: *INDSP_dum* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *ABSACCR* is the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *PERSIST* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise; *FFOLLOW* is the average of the number of firms covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *IFOLLOW* is the average number of two-digit SIC industries covered, during the long-horizon forecast window in year *t*, by each analyst who issues a forecast for firm *j* during that window; *DIFANQ* is the level of the quality proxy for the 'best' analyst minus the level of the quality proxy for the 'worst' analyst, where the 'worst' and 'best' analysts are determined according to various analyst quality proxies; *ANQ_B* is the level of the quality proxy for the 'best' analyst, where the 'best' analyst is determined according to various analyst quality proxies; *HORIZON_B* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst, where 'best' analyst is determined according to various analyst quality proxies; *DIFHORIZON* is the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'worst' analyst minus the number of days between forecast estimation/revision date and subsequent actual earnings reporting date of the 'best' analyst , where the 'worst' and 'best' analysts are determined according to various analyst quality proxies; *YEAR* is the indicator variable for each year 1989–2010. *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

Appendix E: First Stage Regressions of Two-Stage Least Squares (2SLS)

Table E.1: First Stage of Two-Stage Least Squares (2SLS)

Dependent Variable: *INDSP_cont*					
Columns		(1)	(2)	(3)	(4)
		Tests of H1a		Tests of H1b	
	Pred.	BCK Model	Payne Model	BCK Model	Payne Model
INDRELSIZE	+	1.3923***	−0.0191***	0.9041***	0.8585***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
CYCLE	+	0.0002***	−0.0000*	0.0005***	0.0004***
		(<0.001)	(0.097)	(<0.001)	(<0.001)
DISP	?	−0.0213	0.5561***	0.0064	−0.0146
		(0.288)	(<0.001)	(0.763)	(0.504)
HORIZON	?	0.0012***		0.0054**	
		(<0.001)		(0.049)	
SIZE	+	0.0007***	0.0003***	0.0042***	0.0039***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
NUMEST	?	0.0002	−0.0002***	−0.0069***	−0.0002***
		(0.585)	(<0.001)	(<0.001)	(<0.001)
ZSCORE	?	0.0019***		0.0035***	
		(<0.001)		(0.001)	
LOSS	−	−0.0004	0.0056***	−0.0269***	0.0112***
		(0.658)	(<0.001)	(<0.001)	(<0.001)
ABSECHG	−	−0.0323***	0.1089***	0.0024***	−0.0418***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	0.0029		0.0126***	
		(0.256)		(<0.001)	
EL	−	−0.0004		−0.0012***	
		(0.113)		(<0.001)	
ABSACCR	−		−0.0019		−0.0022
			(0.174)		(0.564)
PERSIST	−		−0.0006**		−0.0026***
			(0.016)		(<0.001)
YEAR		yes	yes	yes	yes
CONSTANT		−0.0046**	−0.0081***	−0.0124	−0.0019
		(0.016)	(<0.001)	(0.435)	(0.366)
N		28,701	31,358	23,131	23,422
R^2		0.414	0.310	0.209	0.193

P-values of the coefficients are two-tailed reported parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results for the first-stage regressions of 2SLS. Columns 1 and 2 report the results for tests of H1a based on the BCK-type model and Payne-type model. Columns 3 and 4 report the results for tests of H1b based on the BCK-type model and Payne-type model.

Variable Definition: *INDSP_cont* is the continuous measure of portfolio-share audit firm industry specialisation, measured as per Equation (8); *INDRELSIZE* is the sum of the total assets of all firms in an industry at the end of the current year divided by the sum of total assets of all firms at the end of the current year; *CYCLE* is the industry-year adjusted length of operating cycle in days; I adjust this measure for industry effects by subtracting the industry-year mean operating cycle from each observation; *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *ABSACCR* is the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *PERSIST* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the annual earnings change, 0 otherwise; *YEAR* is the indicator variable for each year 1989–2010.

Appendix F: Long-Horizon Forecast Errors against Prior Year Audit Firm Industry Specialisation and Prior Year Controls

Table F.1: Long-Horizon Forecast Errors against Prior Year Audit Firm Industry Specialisation and Prior Year Controls
Panel A BCK's Model

Dependent Variable: *ABSFE*					
Columns		(1)	(2)	(3)	(4)
	Pred.	BCK1a	BCK2a	BCK1b	BCK2b
INDSP	–	–0.0345***	–0.0316***	–0.0054***	–0.0053***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
DISP_lag	+		0.2564***		0.2568***
			(<0.001)		(<0.001)
HORIZON	+	0.0202***	0.0216***	0.0201***	0.0216***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
SIZE	?	–0.0087***	–0.0080***	–0.0086***	–0.0078***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
NUMEST	–	0.0055***	0.0053***	0.0056***	0.0054***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
ZSCORE_lag	+	0.0022***	0.0015***	0.0021***	0.0014***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
LOSS	+	0.0657***	0.0617***	0.0659***	0.0620***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG_lag	+	0.0973***	0.0707***	0.0964***	0.0693***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
STDROE	+	0.0033	0.0071	0.0068	0.0109
		(0.656)	(0.363)	(0.353)	(0.168)
EL	+	0.0002	0.0002	–0.0000	–0.0000
		(0.660)	(0.699)	(0.985)	(0.944)
YEAR		yes	yes	yes	yes
CONSTANT		–0.0268	–0.0436	–0.0273	–0.0444
		(0.296)	(0.175)	(0.288)	(0.168)
N		26,679	22,217	26,679	22,217
R^2		0.294	0.292	0.295	0.293

P-values of the coefficients are two-tailed reported parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results of the long-horizon regressions based on Models BCK 1a, 1b, 2a and 2b, re-estimated using the one-year lagged values of the control variables that depend on the realisation of the earnings being forecast (*DISP*, *ZSCORE* and *ABSECHG*).

Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation (all models suffixed 'a'), measured as per Equation (8) (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), measured as per Equation (9); *DISP_lag* is the lagged value of the forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE_lag* is the lagged value of Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG_lag* is the lagged value of the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1989–2010.

Panel B Payne's Model

Dependent Variable: *ABSFE*					
Columns		(1)	(2)	(3)	(4)
	Pred..	Payne 1a	Payne 2a	Payne 1b	Payne 2b
INDSP	–	–0.3078**	–0.0258**	–0.0230	–0.0063***
		(0.026)	(0.012)	(0.132)	(<0.001)
DISP_lag	+	0.0183	0.1608***	0.0197	0.1817***
		(0.128)	(0.005)	(0.102)	(0.002)
SIZE	?	0.0557***	0.0009***	0.0556***	0.0005**
		(<0.001)	(0.004)	(<0.001)	(0.041)
NUMEST	–	–0.0072***	–0.0008***	–0.0072***	–0.0008***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSACCR_lag	+	0.1089**	–0.0012	0.1083**	0.0039
		(0.048)	(0.843)	(0.050)	(0.510)
LOSS	+	0.3593***	0.0592***	0.3589***	0.0566***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
ABSECHG_lag	+	0.4018***	0.0523***	0.3841***	0.0636***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
PERSIST_lag	–	–0.1169***	–0.0176***	–0.1164***	–0.0173***
		(<0.001)	(<0.001)	(<0.001)	(<0.001)
YEAR		yes	yes	yes	yes
INDUSTRY		yes	yes	yes	yes
CONSTANT		0.3307**	0.0875***	0.2601*	0.0496***
		(0.049)	(<0.001)	(0.072)	(<0.001)
N		19,437	19,564	19,467	19,564
R^2		0.174	0.296	0.174	0.280

P-values of the coefficients are two-tailed reported parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results of the long-horizon regressions based on Models Payne 1a, 1b, 2a and 2b, re-estimated using the one-year lagged values of the control variables that depend on the realisation of the earnings being forecast (*DISP*, *ABSACCR*, *ABSECHG* and *PERSIST*).

Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *INDSP* is the continuous measure of portfolio-share audit firm industry specialisation (all models suffixed 'a'), measured as per Equation (8) (*INDSP_cont*); or the dichotomous measure of portfolio-share audit firm industry specialisation (all models suffixed 'b'), measured as per Equation (9); *DISP_lag* is forecast dispersion, measured as the lagged value of the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price (deflated by the absolute value of the mean EPS forecast during the period in Models Payne 1a and 1b); *SIZE* is the natural log of total assets; *NUMEST* is the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ABSACCR_lag* is the lagged value of the absolute total accruals, measured as the difference between net income and cash flow from operations divided by lagged total assets; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG_lag* is the lagged value of the absolute value of the change in annual earnings, deflated by beginning-of-month stock price (deflated by the natural log of total assets in Models Payne 1a and 1b); *PERSIST_lag* equals 1 if observation lies between the 20th and 80th percentiles of distribution of the prior year annual earnings change, 0 otherwise; *YEAR* is the indicator variable for each year 1989–2010; *INDUSTRY* is the indicator variable representing two-digit SIC code numbers.

Appendix G: Estimation of Innate Accrual Quality Models

The McNichols's (2002) modification of the Dechow and Dichev's (2002) accrual quality model regress measures accrual quality as the standard deviation of the residuals from firm-specific regressions of working capital accruals on last year, current and one-year-ahead cash flow from operations, change in current sales and the level of current property, plant and equipment, as per Equation (14).

Modified DD accrual quality (MDAQ):

$$\Delta WC_t = \beta_0 + \beta_1 CFO_{t-1} + \beta_2 CFO_t + \beta_3 CFO_{t+1} + \beta_4 \Delta Sales + \beta_5 PPE + \varepsilon \qquad (14)$$

Where

ΔWC_t	=	the change in working capital, computed as change in accounts receivables + change in inventory – change in accounts payable – taxes payable + change in other assets;
CFO_{t-1}	=	Cash flow from operations, last period;
CFO_t	=	Cash flow from operations, current period;
CFO_{t+1}	=	Cash flow from operations, next period;
$\Delta Sales$	=	the change in sales; and
PPE	=	Property, plant and equipment.

The residuals from the above regressions reflect the accruals that are uncorrelated with cash flow realisations. The standard deviations of these residuals across five-year periods are computed and are used as a measure of accrual quality, a larger value of which indicates poor accrual quality.

Francis et al. (2005a) model innate accrual quality as a function of firm size, the standard deviation of cash flow from operations, the standard deviation of sales revenues, firm's operating cycle and the incidence of negative earnings, as per Equation (15).

Innate accrual quality:

$$INNATEAQ_t = \beta_0 + \beta_1 SIZE_t + \beta_2 STDCF_t + \beta_3 STDSALE_t + \beta_4 CYCLE_t + \beta_5 NEARN_t + \varepsilon \qquad (15)$$

Where

$INNATEAQ_t$	=	Modified Dechow and Dichev accrual quality measure, which is the standard deviations of the residuals from Equation (14) calculated over years *t-4* through *t;*
$SIZE$	=	the natural log of total assets;
$STDCF_t$	=	the standard deviation of cash flow from operations;
$STDSALE_t$	=	the standard deviation of revenues;
$CYCLE_t$	=	the length of firm's operating cycle, measured as the natural log of the sum of days accounts receivable and days inventory; and
$NEARN_t$	=	the incidence of negative earnings realisation, measured as the number of years, out of the past 10, where a firm reported a negative earnings.

The fitted value from Equation (15) is the innate accrual quality measure, a larger value of which indicates poorer innate accrual quality.

Appendix H: Long-Horizon Forecast Errors against Prior Year Audit Firm Industry Specialisation for Years 1993–2010

Table H.1: Long-Horizon Forecast Errors against Prior Year Audit Firm Industry Specialisation for Year 1993–2010

Dependent Variable: *ABSFE*			
Columns		(1)	(2)
	Pred.	Full Sample (1993–2010)	Matched Sample (1993–2010)
INDSP	–	–0.0018**	–0.0024**
		(0.035)	(0.026)
DISP	+	0.3935***	0.4243***
		(<0.001)	(0.007)
HORIZON	+	0.0224***	0.0329***
		(<0.001)	(<0.001)
SIZE	?	0.0003	0.0001
		(0.263)	(0.817)
NUMEST	–	–0.0011**	0.0007
		(0.011)	(0.242)
ZSCORE	+	0.0010***	0.0010*
		(0.001)	(0.075)
LOSS	+	0.0195***	0.0210***
		(<0.001)	(<0.001)
ABSECHG	+	0.4929***	0.5028***
		(<0.001)	(<0.001)
STDROE	+	–0.0196***	–0.0175
		(<0.001)	(0.178)
EL	+	–0.0015***	–0.0019***
		(<0.001)	(0.008)
YEAR		yes	yes
INDUSTRY		yes	yes
CONSTANT		–0.1321***	–0.1723***
		(<0.001)	(<0.001)
N		22,742	3,642
R^2		0.564	0.584

Robust *p*-values of the coefficients for *INDSP* are one-tailed reported and for others are two-tailed reported in parentheses (*** $p<0.01$, ** $p<0.05$, * $p<0.1$).

This table presents the results of the long-horizon model to test H1b over a shorter sample period.
Variable Definitions: *ABSFE* is analysts' absolute earnings forecast errors, measured as per Equation (6b); *INDSP* is the dichotomous measure of portfolio-share audit firm industry specialisation, measured as per Equation (9); *DISP* is forecast dispersion, measured as the standard deviation of analysts' forecast EPS deflated by the beginning-of-month stock price; *HORIZON* is the natural log of the average number of days between mean forecast estimation date and subsequent actual earnings reporting date; *SIZE* is the natural log of the market value of equity; *NUMEST* is the natural log of the number of analysts issuing earnings forecasts for the firm in the 90-day window prior to earnings reporting; *ZSCORE* is Zmijewski's financial distress score; *LOSS* equals 1 if a firm reports negative earnings, 0 otherwise; *ABSECHG* is the absolute value of the change in annual earnings, deflated by beginning-of-month stock price; *STDROE* is the standard deviation of return on equity over the previous five years; *EL* is earnings per share, winsorized at 5 (-5); *YEAR* is the indicator variable for each year 1993–2010.